

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5

(i) APPLICANT:

- (A) NAME: COMMONWEALTH SCIENTIFIC AND INDUSTRIAL
RESEARCH ORGANISATION
(B) STREET: Limestone Avenue
(C) CITY: Campbell
(D) STATE: ACT
(E) COUNTRY: AUSTRALIA
(F) POSTAL CODE (ZIP): 2612

10

- (A) NAME: THE AUSTRALIAN NATIONAL UNIVERSITY
(B) STREET: BRIAN LEWIS CRESCENT
(C) CITY: ACTON
(D) STATE: ACT
(E) COUNTRY: AUSTRALIA
(F) POSTAL CODE (ZIP): 2601

15

- (A) NAME: GOODMAN FIELDER LIMITED
(B) STREET: LEVEL 42, GROSVENOR PLACE
(C) CITY: SYDNEY
(D) STATE: NSW

20

- (E) COUNTRY: AUSTRALIA
(F) POSTAL CODE (ZIP): 2000

25

- (A) NAME: GROUPE LIMAGRAN PACIFIC PTY LIMITED
(B) STREET: LEVEL 31, 1 O'CONNELL STREET
(C) CITY: SYDNEY
(D) STATE: NSW

30

- (E) COUNTRY: AUSTRALIA
(F) POSTAL CODE (ZIP): 2000

35

(ii) TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS

(iii) NUMBER OF SEQUENCES: 17

40

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

45

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "pcr primer based on the N-terminal sequence of wSBE I 5'-end at position 168 of SEQ ID NO:5"

55

(iii) HYPOTHETICAL: NO

PCR
primer
X(D)

- 62 -

(iv) ANTI-SENSE:

(v) FRAGMENT TYPE:

5 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: triticum tauschii
(F) TISSUE TYPE: Endosperm

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGCACCGCGAG AGACTGG

17

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "pcr primer in which 5' end is at position 1590 of SEQ ID NO:5"

(iii) HYPOTHETICAL: NO

25 (iv) ANTI-SENSE:

(v) FRAGMENT TYPE:

30 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: triticum tauschii
(F) TISSUE TYPE: Endosperm

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

35 TACATTTCCCT TGTCCATCA

19

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

45 (A) DESCRIPTION: /desc = "pcr primer 5' end is at position 1 of SEQ ID NO:5"

(iii) HYPOTHETICAL: NO

50 (iv) ANTI-SENSE:

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- 55 (A) ORGANISM: triticum tauschii
(F) TISSUE TYPE: Endosperm

- 63 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATCACCGAGAG CTTGCTCA 18

5 (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

10

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "pcr primer 5' end is at position 334 of SEQ ID NO:5"

15 (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE:

20 (v) FRAGMENT TYPE:

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: triticum tauschii
 - (F) TISSUE TYPE: Endosperm

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CGGTACACAG TTGCGTCATT TTC 23

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2687 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: cDNA

- (iii) HYPOTHETICAL: NO

40 (iv) ANTI-SENSE:

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: triticum tauschii
 - (F) TISSUE TYPE: Endosperm

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ATCGACGAAG ATGCTCTGCC TCACCGCCCC CTCCTGCTCG CCATCTCTCC CGCCGCGCCC 60

50 CTCCCGTCCC GCTGCTGACC GGCCCGGACC GGGGATTTCG GCCAAGAGCA AGTTCTCTGT 120

TCCCGTGTCT GCGCCAAGAG ACTACACCAT GGCAACAGCT GAAGATGGTC TTGGCGACCT 180

TCCGATATACT GATCTGGATC CGAACGTTGC CGGCTTCAAG GAACACTTCA GTTATAGGAT 240

55 GAAAAAGTAC CTTGACCAGA AACATTCGAT TGAGAAGCAC GAGGGAGGCC TTGAAGAGTT 300

CTCTAAAGGC TATTTGAAGT TTGGGATCAA CACAGAAAAT GACGCAACTG TGTACCGGGA 360

- 64 -

| | | | | | | | |
|----|-------------|------------|------------|-------------|-------------|------------|------|
| | ATGGGCCCCCT | GCAGCAATGG | ATGCACAACT | TATTGGTGAC | TTCAACAACT | GGATGGCTC | 420 |
| 5 | TGGGCACAGG | ATGACAAAGG | ATAATTATGG | TGTTTGGTCA | ATCAGGATT | CCCATGTCAA | 480 |
| | TGGGAAACCT | GCCATCCCC | ATAATTCAA | GGTTAAATT | CGATTCACC | GTGGAGATGG | 540 |
| | ACTATGGTC | GATCGGGTTC | CTGCATGGAT | TCGTTATGCA | ACTTTGACG | CCTCTAAATT | 600 |
| 10 | TGGAGCTCCA | TATGACGGTG | TTCACTGGG | TCCACCTTCT | GGTGAAAGGT | ATGTGTTAA | 660 |
| | GCATCCTCGG | CCTCGAAAGC | CTGACGCTCC | ACGTATTAC | GAGGCTCATG | TGGGGATGAG | 720 |
| 15 | TGGTGAGAGG | CCTGAAGTAA | GCACATACAG | AGAATTGCA | GACAATGTGT | TACCGCGCAT | 780 |
| | AAAGGCAAAC | AACTACAACA | CAGTCAGCT | GATGGCAATC | ATGGAACATT | CCATATTATG | 840 |
| | CTTCTTTGG | TACCATGTGA | CGAATTCTT | CGCAGTTAGC | AGCAGATCAG | GAACACCAGA | 900 |
| 20 | GGACCTCAAA | TATCTGTTG | ACAAGGCACA | TAGCTTAGGG | TTGCGTGTTC | TGATGGATGT | 960 |
| | TGTCCATAGC | CATGCGAGCA | GTAATATGAC | AGATGGTCTA | AATGGCTATG | ATGTTGGACA | 1020 |
| 25 | AAACACACAG | GAGTCCTATT | TCCATACAGG | AGAAAGGGT | TATCATAAAC | TGTGGATAG | 1080 |
| | TCGCCTGTT | AACTATGCCA | ATTGGGAGGT | CTTACGGTAT | CTTCTTTCTA | ATCTGAGATA | 1140 |
| | TTGGATGGAC | GAATTCATGT | TTGACGGCTT | CCGATTTGAT | GGAGTAACAT | CCATGCTATA | 1200 |
| 30 | TAATCACCAT | GGTATCAATA | TGTCATTGCG | TGGAAATTAC | AAGGAATATT | TTGGTTGGA | 1260 |
| | TACCGATGTA | GATGCAGTTG | TTTACATGAT | GCTTGCAC | CATTAAATGC | ACAAAATCTT | 1320 |
| 35 | GCCAGAAGCA | ACTGTTGTTG | CAGAAGATGT | TTCAGGCATG | CCAGTGCTT | GTCGGTCAGT | 1380 |
| | TGATGAAGGT | GGAGTAGGGT | TTGACTATCG | CCTTGCTATG | GCTATTCTG | ATAGATGGAT | 1440 |
| | TGACTACTTG | AAGAACAAAG | ATGACCTTGA | ATGGTCAATG | AGTGAATAG | CACACTCT | 1500 |
| 40 | GACCAACAGG | AGATATACGG | AAAAGTCAT | TGCATATGCT | GAGGCCACG | ATCAGTCTAT | 1560 |
| | TGTTGGCGAC | AAGACTATGG | CATTCTCTT | GATGGACAAG | GAAATGTATA | CTGGCATGTC | 1620 |
| 45 | AGACTTGCAG | CCTGCTTCAC | CTACAATTGA | TCGTGGAATT | GCACCTCAA | AGATGATTCA | 1680 |
| | CTTCATCACC | ATGGCCCTTG | GAGGTGATGG | CTACTTGAAT | TTTATGGTA | ATGAGTTGG | 1740 |
| | CCACCCAGAA | TGGATTGACT | TTCCAAGAGA | AGGCAACAAAC | TGGAGTTATG | ATAAATGCAG | 1800 |
| 50 | ACGCCAGTGG | AGCCTCTCAG | ACATTGATCA | CCTACGATAC | AAGTACATGA | ACGCATTGGA | 1860 |
| | TCAAGCAATG | AATGCGCTCG | ACGACAAGTT | TTCCTTCCTA | TCGTCAATCAA | AGCAGATTGT | 1920 |
| 55 | CAGCGACATG | AATGAGGAAA | AGAAGATTAT | TGTATTGAA | CGTGGAGATC | TGGTCTTCGT | 1980 |
| | CTTCAATT | CATCCCAGTA | AAACTTATGA | TGGTTACAAA | GTCGGATGTG | ATTTGCCTGG | 2040 |
| | GAAGTACAAG | GTAGCTCTGG | ACTCCGATGC | TCTGATGTT | GGTGGACATG | GAAGACTGGC | 2100 |
| 60 | CCAGTACAAC | GATCACTTCA | CGTCACCTGA | AGGAGTACCA | GGAGTACCTG | AAACAAACTT | 2160 |
| | CAACAACCGC | CCTAATTCA | TCAAAGTCCT | GTCTCCACCC | CGCACTTGTG | TGGCTTACTA | 2220 |
| 65 | TCGCGTCGAG | AAAAAGCGG | AAAAGCTAA | GGATGAAGGA | GCTGCTTCTT | GGGGCAAAGC | 2280 |
| | TGCTCCTGGG | TACATCGATG | TTGAAGCCAC | TCGTGTCAA | GACGCAGCAG | ATGGTGAGGC | 2340 |

- 65 -

GACTTCTGGT TCCAAAAGG CGTCTACAGG AGGTGACTCC AGCAAGAAGG GAATTAACCTT 2400
 TGTCTTCGGG TCACCTGACA AAGATAACAA ATAAGCACCA TATCAACGCT TGATCAGAAC 2460
 5 CGTGTACCGA CGTCCTTGTA ATATTCTGC TATTGCTAGT AGTAGCAATA CTGTCAAAC 2520
 GTGCAGACTT GAGATTCTGG CTTGGACTTT GCTGAGGTTA CCTACTATAT AGAAAGATAA 2580
 10 ATAAGAGGTG ATGGTGCGGG TCGAGTCCGG CTATATGTGC CAAATATGCG CCATCCCGAG 2640
 TCCTCTGTCA TAAAGGAAGT TTCGGGCTTT CAGCCCAGAA TAAAAAA 2687

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 807 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- 20 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- 25 (iv) ANTI-SENSE:
- (v) ORIGINAL SOURCE:
 - (A) ORGANISM: triticum tauschii
 - (F) TISSUE TYPE: Endosperm
- 30 (ix) FEATURE:
 - (A) NAME/KEY: Protein
 - (B) LOCATION: 1..807
 - (D) OTHER INFORMATION:/label= sbeI
 /note= "deduced amino acid sequence from SEQ ID NO:5"
- 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

| | | | | | | | | | | | | | | | | |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | Met | Leu | Cys | Leu | Thr | Ala | Pro | Ser | Cys | Ser | Pro | Ser | Leu | Pro | Pro | Arg |
| 40 | 1 | | | | | | | | 5 | | 10 | | | | 15 | |
| | Pro | Ser | Arg | Pro | Ala | Ala | Asp | Arg | Pro | Gly | Pro | Gly | Ile | Ser | Ala | Lys |
| | | | | | | | | 20 | | 25 | | | | 30 | | |
| 45 | Ser | Lys | Phe | Ser | Val | Pro | Val | Ser | Ala | Pro | Arg | Asp | Tyr | Thr | Met | Ala |
| | | | | | | | | 35 | | 40 | | | | 45 | | |
| | Thr | Ala | Glu | Asp | Gly | Val | Gly | Asp | Leu | Pro | Ile | Tyr | Asp | Leu | Asp | Pro |
| 50 | | | | | | | | 50 | | 55 | | | | 60 | | |
| | Lys | Phe | Ala | Gly | Phe | Lys | Glu | His | Phe | Ser | Tyr | Arg | Met | Lys | Lys | Tyr |
| | | | | | | | | 65 | | 70 | | | | 75 | | 80 |
| 55 | Leu | Asp | Gln | Lys | His | Ser | Ile | Glu | Lys | His | Glu | Gly | Gly | Leu | Glu | Glu |
| | | | | | | | | 85 | | 90 | | | | 95 | | |
| | Phe | Ser | Lys | Gly | Tyr | Leu | Lys | Phe | Gly | Ile | Asn | Thr | Glu | Asn | Asp | Ala |
| | | | | | | | | 100 | | 105 | | | | 110 | | |
| 60 | Thr | Val | Tyr | Arg | Glu | Trp | Ala | Pro | Ala | Ala | Met | Asp | Ala | Gln | Leu | Ile |
| | | | | | | | | 115 | | 120 | | | | 125 | | |

- 66 -

| | |
|----|-----------------------------------------------------------------|
| | Gly Asp Phe Asn Asn Trp Asn Gly Ser Gly His Arg Met Thr Lys Asp |
| | 130 135 140 |
| 5 | Asn Tyr Gly Val Trp Ser Ile Arg Ile Ser His Val Asn Gly Lys Pro |
| | 145 150 155 160 |
| | Ala Ile Pro His Asn Ser Lys Val Lys Phe Arg Phe His Arg Gly Asp |
| | 165 170 175 |
| 10 | Gly Leu Trp Val Asp Arg Val Pro Ala Trp Ile Arg Tyr Ala Thr Phe |
| | 180 185 190 |
| | Asp Ala Ser Lys Phe Gly Ala Pro Tyr Asp Gly Val His Trp Asp Pro |
| 15 | 195 200 205 |
| | Pro Ser Gly Glu Arg Tyr Val Phe Lys His Pro Arg Pro Arg Lys Pro |
| | 210 215 220 |
| 20 | Asp Ala Pro Arg Ile Tyr Glu Ala His Val Gly Met Ser Gly Glu Arg |
| | 225 230 235 240 |
| | Pro Glu Val Ser Thr Tyr Arg Glu Phe Ala Asp Asn Val Leu Pro Arg |
| | 245 250 255 |
| 25 | Ile Lys Ala Asn Asn Tyr Asn Thr Val Gln Leu Met Ala Ile Met Glu |
| | 260 265 270 |
| | His Ser Ile Leu Cys Phe Phe Trp Tyr His Val Thr Asn Phe Phe Ala |
| 30 | 275 280 285 |
| | Val Ser Ser Arg Ser Gly Thr Pro Glu Asp Leu Lys Tyr Leu Val Asp |
| | 290 295 300 |
| 35 | Lys Ala His Ser Leu Gly Leu Arg Val Leu Met Asp Val Val His Ser |
| | 305 310 315 320 |
| | His Ala Ser Ser Asn Met Thr Asp Gly Leu Asn Gly Tyr Asp Val Gly |
| | 325 330 335 |
| 40 | Gln Asn Thr Gln Glu Ser Tyr Phe His Thr Gly Glu Arg Gly Tyr His |
| | 340 345 350 |
| | Lys Leu Trp Asp Ser Arg Leu Phe Asn Tyr Ala Asn Trp Glu Val Leu |
| 45 | 355 360 365 |
| | Arg Tyr Leu Leu Ser Asn Leu Arg Tyr Trp Met Asp Glu Phe Met Phe |
| | 370 375 380 |
| 50 | Asp Gly Phe Arg Phe Asp Gly Val Thr Ser Met Leu Tyr Asn His His |
| | 385 390 395 400 |
| | Gly Ile Asn Met Ser Phe Ala Gly Asn Tyr Lys Glu Tyr Phe Gly Leu |
| | 405 410 415 |
| 55 | Asp Thr Asp Val Asp Ala Val Val Tyr Met Met Leu Ala Asn His Leu |
| | 420 425 430 |
| | Met His Lys Ile Leu Pro Glu Ala Thr Val Val Ala Glu Asp Val Ser |
| 60 | 435 440 445 |
| | Gly Met Pro Val Leu Cys Arg Ser Val Asp Glu Gly Gly Val Gly Phe |
| | 450 455 460 |

- 67 -

| | | |
|----|-----------------------------------------------------------------|--|
| | Asp Tyr Arg Leu Ala Met Ala Ile Pro Asp Arg Trp Ile Asp Tyr Leu | |
| | 465 470 475 480 | |
| 5 | Lys Asn Lys Asp Asp Leu Glu Trp Ser Met Ser Ala Ile Ala His Thr | |
| | 485 490 495 | |
| | Leu Thr Asn Arg Arg Tyr Thr Glu Lys Cys Ile Ala Tyr Ala Glu Ser | |
| | 500 505 510 | |
| 10 | His Asp Gln Ser Ile Val Gly Asp Lys Thr Met Ala Phe Leu Leu Met | |
| | 515 520 525 | |
| | Asp Lys Glu Met Tyr Thr Gly Met Ser Asp Leu Gln Pro Ala Ser Pro | |
| | 530 535 540 | |
| 15 | Thr Ile Asp Arg Gly Ile Ala Leu Gln Lys Met Ile His Phe Ile Thr | |
| | 545 550 555 560 | |
| 20 | Met Ala Leu Gly Gly Asp Gly Tyr Leu Asn Phe Met Gly Asn Glu Phe | |
| | 565 570 575 | |
| | Gly His Pro Glu Trp Ile Asp Phe Pro Arg Glu Gly Asn Asn Trp Ser | |
| | 580 585 590 | |
| 25 | Tyr Asp Lys Cys Arg Arg Gln Trp Ser Leu Ser Asp Ile Asp His Leu | |
| | 595 600 605 | |
| 30 | Arg Tyr Lys Tyr Met Asn Ala Phe Asp Gln Ala Met Asn Ala Leu Asp | |
| | 610 615 620 | |
| | Asp Lys Phe Ser Phe Leu Ser Ser Ser Lys Gln Ile Val Ser Asp Met | |
| | 625 630 635 640 | |
| 35 | Asn Glu Glu Lys Lys Ile Ile Val Phe Glu Arg Gly Asp Leu Val Phe | |
| | 645 650 655 | |
| | Val Phe Asn Phe His Pro Ser Lys Thr Tyr Asp Gly Tyr Lys Val Gly | |
| | 660 665 670 | |
| 40 | Cys Asp Leu Pro Gly Lys Tyr Lys Val Ala Leu Asp Ser Asp Ala Leu | |
| | 675 680 685 | |
| | Met Phe Gly Gly His Gly Arg Val Ala Gln Tyr Asn Asp His Phe Thr | |
| | 690 695 700 | |
| 45 | Ser Pro Glu Gly Val Pro Gly Val Pro Glu Thr Asn Phe Asn Asn Arg | |
| | 705 710 715 720 | |
| 50 | Pro Asn Ser Phe Lys Val Leu Ser Pro Pro Arg Thr Cys Val Ala Tyr | |
| | 725 730 735 | |
| | Tyr Arg Val Glu Glu Lys Ala Glu Lys Pro Lys Asp Glu Gly Ala Ala | |
| | 740 745 750 | |
| 55 | Ser Trp Gly Lys Ala Ala Pro Gly Tyr Ile Asp Val Glu Ala Thr Arg | |
| | 755 760 765 | |
| | Val Lys Asp Ala Ala Asp Gly Glu Ala Thr Ser Gly Ser Lys Lys Ala | |
| | 770 775 780 | |
| 60 | Ser Thr Gly Gly Asp Ser Ser Lys Lys Gly Ile Asn Phe Val Phe Gly | |
| | 785 790 795 800 | |

Ser Pro Asp Lys Asp Asn Lys
805

- 5 (2) INFORMATION FOR SEQ ID NO: 7:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 319 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - 10 (ii) MOLECULE TYPE: cDNA
 - (iii) HYPOTHETICAL: NO
 - 15 (iv) ANTI-SENSE:
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: triticum tauschii
 - (F) TISSUE TYPE: Endosperm
 - 20 (ix) FEATURE:
 - (A) NAME/KEY: misc_signal
 - (B) LOCATION: 1..319
 - (D) OTHER INFORMATION:/function= "3' untranslated region of wSBE I-D4 cDNA"
 - 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

| | | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|-----|
| GCGACTTCTG | GTTCCAAAAA | GGCGTCTACA | GGAGGTGACT | CCAGCAAGAA | GGGAATTAAC | 60 | |
| 30 | TTTGTCTTCG | GGTCACCTGA | CAAAGATAAC | AAATAAGCAC | CATATCAACG | CTTGATCAGA | 120 |
| | ACCGTGTACC | GACGTCCTTG | TAATATTCT | GCTATTGCTA | GTAGTAGCAA | TACTGTCAAA | 180 |
| 35 | CTGTGCAGAC | TTGAGATTCT | GGCTTGGACT | TTGCTGAGGT | TACCTACTAT | ATAGAAAGAT | 240 |
| | AAATAAGAGG | TGATGGTGGC | GGTCGAGTCC | GGCTATATGT | GCCAAATATG | CGCCATCCCG | 300 |
| 40 | AGTCCTCTGT | CATAAAGGA | | 319 | | | |
 - (2) INFORMATION FOR SEQ ID NO: 8:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4890 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - 45 (ii) MOLECULE TYPE: DNA (genomic)
 - (iii) HYPOTHETICAL: NO
 - (iv) ANTI-SENSE:
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: triticum tauschii
 - (F) TISSUE TYPE: Endosperm
 - 55 (ix) FEATURE:

(A) NAME/KEY: promoter
 (B) LOCATION: 1..4890
 (D) OTHER INFORMATION:/function= "promoter containing sequence of SBE I"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

| | | | | | | | |
|----|-------------|---------------|------------|-------------|------------|-------------|------|
| | GGGTGGCGGG | TCGGGCGGCA | AGGCGCGGGG | CGGCAGGGCG | GCCGGGGCG | CGCGGGCG | 60 |
| 10 | CGGGCGGCAG | CGGCGGCTAG | GGTTTCGCGG | CGGCAGGCGAC | TTGGGCTGAG | GCGGGGCACG | 120 |
| | GGCTGCGGCT | TTAAAGGCCG | GCCAGGCTGA | GGTGTCCGGG | TCGGACACGG | CCCGTAAGGC | 180 |
| 15 | GGTTGACTTT | AAAAAAATAAT | AATTGGACA | TGCAAAAAAG | TAAGAAAAGA | AATAATAAAC | 240 |
| | GGACTCCAAA | AATCCCGAAG | TAAATTTTC | CCCATTCTTA | AAAATAAGCC | GGACAAGATG | 300 |
| | AACATTATT | TGGGCCTAAA | ATGCAATT | TTT GCG | TATTTTCCT | AATTGGAAT | 360 |
| 20 | AAAATCAAAT | AAAATCCAAA | AAAATCAA | TATTTGTTT | TAATATTTT | CCTCCAATAT | 420 |
| | TTCATTATT | GTGAAGAAGT | CATTTATCC | CATCTCATAT | ATTTTGATAT | GAAATATT | 480 |
| 25 | CGGAGAGAAA | AATAATTAAA | ACAAATGATC | CTATTTCAA | AATTGAGAA | AACCCAAATA | 540 |
| | TGAAAATAAC | GAAATCCCCA | ACTCTCTCCG | TGGGTCCCTG | AGTTGCGTGA | AATTCTAGG | 600 |
| | ATCACAAAATC | AAAATGCAAT | AAAATATGAT | ATGCATGATG | ATCTAATGTA | TAACATTCCA | 660 |
| 30 | ATTGAAAATT | TGGGATGTTA | CATATAACTC | AAATTCTATA | ATTATGAACA | CAGAAATATT | 720 |
| | AATGTAGAAC | TCTATTTGT | TTTGAAATTG | TATTATTTT | TAGAATTAGT | CTAGAGCATT | 780 |
| 35 | TCGTGAAC | TT GAATCAAACC | TTTAAATAAA | ACAAAGCATA | AAAATGACAA | ATTCACATAT | 840 |
| | GAAATAACTT | GTGTTACATA | GATTATTAC | AATAGCGTTG | TATGTGTGTA | TGTGTGCGTG | 900 |
| | AGTGCCTATG | GTAATATCAA | TAAATATCTT | GATAGATGTT | TCTACAATT | ACGGGTCTAA | 960 |
| 40 | CTAGTAATGC | AATGCAATGC | ATGCTAAAG | AATAGAACCT | TAGTTTCATT | TAACTAACAA | 1020 |
| | TTTTCAAATG | TATGAGTTGC | CAACAAGTGG | CATACTGGC | ACTGTTGTT | TGTTCATTT | 1080 |
| 45 | ATGGAAAGTT | CTTCTCTTT | TACATGGTT | AGATTCCAGC | ATGTAGCCAC | AAAATATGAT | 1140 |
| | TGTCAAAGA | TAATACCTCA | TAATACAATT | CCACTAAAGT | CACCTAGCCC | AA GTGACCGA | 1200 |
| | CCTGATCCTG | AAATAAAATC | AGAAGATTG | GTGTCATCAT | CATGACAACA | AATTATTAGG | 1260 |
| 50 | CGGTAGATCT | TGTGGTAGTA | CTCATGATGT | AAAATTATCA | AGAGGGAGAG | AATGTATGGA | 1320 |
| | GATTTATGTG | AA GTACATCG | TACACCAGAC | ATAGTTGACA | CATCGATTT | TTAAGATACA | 1380 |
| 55 | TTTGGACGCG | CCTTGTGGGA | GTGAAAGTA | CTACCATGTA | TTAGAAGAGG | TGAAATGAGA | 1440 |
| | AATGCCATAG | CTAGCAAGTA | GGCCTAGTTA | AGGAAATTCT | TCCTTAGATC | CCCTTCTCCC | 1500 |
| | GAAGAGTGAA | GTGCTTCAAC | TAAAGGTTAG | ACCCACTTAA | AAAATGTCAC | TTTGAATCTT | 1560 |
| 60 | TGCTTCCCTT | GTCGTAATCC | TGTGCATTG | TAGGTCCCTC | GGATCTGAGC | CCTTTCTCCA | 1620 |
| | AGCCCTTCAT | TGGATTCCCC | TGGATGTCTT | TTTGTACAT | TTTATTGAAG | TGAGAGTGAA | 1680 |
| 65 | TTATTATATG | CCCATAGGAG | GTGGGATATA | AAGGCTGTTG | GTATTCTGCA | CCATACATGC | 1740 |

- 70 -

TAGAGTAGGG AGGAGAGGCT GGTGCATGAT ACATGGTGG A CTAGCCCATA TATTACCC 1800
 TCCCCCACCC ACTAACAAAGT TTTTTTTATT AGGTCTTCAT CCTCTGATTT GTTTTCTGT 1860
 5 TAGCCCATTC TTCATCATGG ACTTATTAAT CATGATTAGT TTCTTGGATT TTTGTTACT 1920
 TGACTTGAAT TTGACATGT GCCTCATATA TGGCATGTGG GACTGATAGG AAGATATATT 1980
 10 CTCACAAACAT TAACTTAAA AGGATTATTT TTTGGTGCA GTCGTAAAGA AAACACTTT 2040
 CTTTATGCT AAAAGTTATT CAAACATAGA TTTATAAAC AAGGATATCA CCATGCATGA 2100
 CCATGCGCTC TCTCATGTT ACTCTAGAAA CCATATATCT CTTGTTGCA AAATATTTAA 2160
 15 TCTATCCTCC TTGTTCTGG GAATGAGTCG GGGAAAGGTA TCTTAGGGAA GTTAAAGTG 2220
 AGGCAAGTAA GAGCAACTCT AGCAGAGTCG CGATATGCC AATGCCATA ATGCCAATAT 2280
 20 GCCATTTTG GCCCAAATG GCACTTCAGA AGAGTCACCA TATCCCTTCG GATAGCCATA 2340
 ATTTAGGGAG CTCGCTCCAC AAACAAGCTT CGAGCCTCCA AATATGGAGG CCATGGATTC 2400
 GTTGTGCG ACTCACTCCA TATCCAACCG CAAGCGCATG CATGAGGGAA GTTTAGCTT 2460
 25 CTTCCCTCCTT GCGCCAACGC CGGGATTTA CACAGCGCAT TACAGGTACA TGAACCAGCA 2520
 TGCACAGATA ATCACCGACG AGTGGGGTGA CAAGAAGGAT AAGCACCCCTC CCATTAGTGG 2580
 30 TGCGCCCAC TCCCCTCAAAT TCATGAGGCA GCCATTGGA TGGTCATCGC GTGGCATAAG 2640
 CTCCGACTAT AAAATCTCAA CGGCATCACC AAAACCATAG CTGCCGCCTC CCCCTTCCTC 2700
 GGCAATCACCT CCCCAAGACA TCTCCTCCCC TCTATGCCAC AATGTATCA TTATGGAGAG 2760
 35 ACACAACATAC TGGTAAACCG CATAACCAAT CATGGTTAC CGGCAGTGCG AACCCACCT 2820
 TCCTCCCACG ATGGTAGGAT ATTCTCCTCC TAGAATGGCG CGTGTGGCGC TTCCCTCCTCC 2880
 40 CGAGGCTGAT ATGTCGGCTC CCATGATGGC GTGCATCATT GATTTGGCGC TTCGGGTCCA 2940
 TCATACATGT TAACGAGGTC ATCCCCATTG ATGTCGTTGG TCCCCTTGCC CCCCAGTCGG 3000
 ATCCTGAGGA CCCGTTCGAT GTCGCAATGC GACTCTCAA ACTCAAAGCT CACAATGAGG 3060
 45 AGTACGTCCT CTAGGAGTTG CGCCCCGCAA CCATCTATAA GGAGGAGCAA CGATAGCTCT 3120
 CCCCTACGCC TTCCTCGACG ATCTCTCTTA GGAGGACAAC GGCTAGACGA CGCGCGCGC 3180
 50 GGCGAAGGTA CTGCAGGTAG TAGAACATAG CAATGTCGAA TGGCGACATT GCATATTTG 3240
 AAAATGTCGC TCAACGACTT TTGAAGTCGC AAATAAAATG TAGTGTGACT ACTTTGGCC 3300
 AGCAATATAA GTTTATCACA TTTGATAATG ATTTGAACCG GTGTGGTCA ACTAAATGTA 3360
 55 CCATAAAATTG AACATACAAA TTTTAGCAA ATGAAAAAAG AAACAAGTAA GACCACAAAT 3420
 ATGAAAGCCG CATATCGCGA CTATGTGTTT GAGCCGCAGC TGCCAAAGTAC ATATGAAGCG 3480
 60 TACTCCATAT GACATACGAC AACCATACAT ATGAAGACTC TACTAGAGTT CTCTAAGGCC 3540
 GCTTTAGCG CCTTTCGTGC AGTGGTGCCTC ATAGGGAGTG AGGGTAGTTG GACTGTTCGT 3600
 TTCCCTTTT TTCATTTCTT TGAAATCTAT TTTATTTTT TTCTCTTTG TAGGTTCCC 3660
 65 AAATTTATAT ACCATTTTC TGTTCTCGC TATTTTTGT TGTTATATTC TAGTTTCATA 3720
 TTTTCTATT ATTAATTTGT GTCTCTTATG AGAAGTCCAG ACTTGCATAT GGAGGTGCAC 3780

- 71 -

ACACAAACAT ATAAAGTATA AATACTAATC TGAGAAGTAT GTTTGCCTGG TCAAAAAAAC 3840
5 ATCATCAAAA CCTGCCAATA TGAGATATAG TTTTGAAATAT ATCAATATGA GCAACGCAAC 3900
CATTTAAAAT GTGAACAATT GTTTTTTAG AAAAAATATA AGAAATAACT CCAACCCAGC 3960
CAAACCACAT GCTATACACT TGCTCCATAT GAAACCATGT TTGCTATTGG GCAGTTGCCT 4020
10 GAAACCGAAA GTAATGTTAG CCGTTTTCT ATTCAAAGAA GAAGGAGAGT CGAGGTGACG 4080
CGATGCTTAG ACGTGAGATG GGGATGACCA CAACGTCCT ACAGAGACCT CACCGGAGAT 4140
15 GGGGACATTG CAGTTGACAC GAGAGCGGTG AGGGGCTGCG ATGCGTGTGC GGCAACATGT 4200
GGCGAGGCAG ACGTCGGGCT GGCAGGTAGG GGGGAGGGGG AAGGACCGGG GGAGGAAGAA 4260
GAGGAGTAGC CTGCAAAACA TGGTACACCA GTTTCTGCC CTACGAAAAC CTCATTCAT 4320
20 TCCCCCACCC TGACAAGCAA CAACCAACCA TCGCAGTCCC ACATGTCCT CTGGTCTTG 4380
CAAAAAGTAA TTGTTCTTGC TGGACAGCGC AAAGAGTAAA CTTTGTTAG TTTTCATTTC 4440
25 TAGAAAAAGC AATCCTTTA TAGTTCTTT GTGAAAGTAA TGCTTTATA GTGATTGGG 4500
TGTTCTTTA GAGCAAATAT CTTCTTTTT TTTTAGGGAA AAGAGCAAAT ATCTTCCACT 4560
TTTCACAAAAA CTGACGAAGG CTGAAAGTGG CGAGACAGTG AGGGCCATA GCTTCGTCC 4620
30 GGCCCAGCGG CGCACGACCG TCCACGTGCA CCCCGCCCT CCCGGGCCCG CAGATCCGTT 4680
CTCCCTCGCC CCCGTTCCC CCTCCCTCCC TCTCGTTGCT TCCACTCCAC TGTTCTCCTC 4740
35 TTCCTGTCCA AAGCGGCCAC GGACCGGAAA AAAATCACGC CTTTCCGTTG GGTCTCCGGC 4800
GCCACACTCC TCCTCCGGCC GATATAAAGC GCGGGGGCC ACGGGCCCGG CGCAAAATGG 4860
GATTCCCGTC CGCCGCCATG GAGGAAGATG 4890
40 (2) INFORMATION FOR SEQ ID NO: 9:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6228 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
45 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO
50
(iv) ANTI-SENSE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: triticum tauschii
55 (F) TISSUE TYPE: Endosperm

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1
60 (D) OTHER INFORMATION:/product= "coding region of wSBE I-D4 gene"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

- 72 -

| | | |
|----|--------------------------------------------------------------------|------|
| | ACGGGCCCCGG CGCAAAATGG GATTCCCGTC CGCCGCCATC GACGAAGATG CTCTGCCTCA | 60 |
| | CCGCCCCCTC CTGCTGCCA TCTCTCCCGC CGCGCCCCCTC CCGTCCCGCT GCTGACCGGC | 120 |
| 5 | CCGGACCGGG GATCTCGGTG AGTCAGTCGG GATCTTCATT TCTTTCTTT TCTTTCGTTT | 180 |
| | CCGGCTCCGT TCTGCCGGGG TTTCCTGAT GCGATGCCGC GCGCCCGCAG GGCGGCGGCA | 240 |
| 10 | ATGTGCGGCT GAGCGCGGTG CCCGCGCCCT CTTCGCTCCG CTGGTCGTGG CCGCGGAAGG | 300 |
| | TGAGCCCTCT CCCCTGTCTA CCCAGATTG CGACCGTGAT CCCCTGTTGT CGCCGGGCAA | 360 |
| | ACGGAATCTG ATCCACGGTG GTTATTGGAA ATAGTATATA CTACTAATAA ACTTGAGGCT | 420 |
| 15 | GGGATTGTC CACTGAGGAA CAAGTGGATG CGATTCGAT TGGATTTCTC TGCTTTATGC | 480 |
| | GATCCGTACG CAGAATATCC CTCCTGCAGT GTCTAACCG TATTACTGGA TGTACAACCC | 540 |
| 20 | AAATGTGTAT AATCTGTGCT GAATGTATCA ACCAATAATT GCTGCATTGT GAAAACATAA | 600 |
| | TCCTGTGTTG TGTCTCTACT ACTTGTTCAAG TCCCTGATCTG CCGCTTATCC TAACTTTGT | 660 |
| | TCATTATGG AAGGCCAAGA GCAAGTTCTC TGTTCCCGTG TCTGCGCCAA GAGACTACAC | 720 |
| 25 | CATGGCAACA GCTGAAGATG GTGTTGGCGA CCTTCCGATA TACGATCTGG ATCCGAAGTT | 780 |
| | TGCCGGCTTC AAGGAACACT TCAGTTATAG GATGAAAAAG TACCTTGACC AGAAACATTC | 840 |
| 30 | GATTGAGAAG CACGAGGGAG GCCTTGAAGA GTTCTCTAAA GGTTAGCTTT TGTTTCATGT | 900 |
| | GTGAAACA ATAGTTACAT CTTGTGGCGT CCGCAGCACAA AAAGACATAA TGCGACTCTG | 960 |
| | TTTTGTAGGC TATTGAAAGT TTGGGATCAA CACAGAAAAT GACGCAACTG TGTACCGGGA | 1020 |
| 35 | ATGGGCCCCCT GCAGCAATGT AAGTTCTAGT GTTGTACGC AACTAATTGC AATGGTCGTT | 1080 |
| | GGTTAACTTA TGAAGTGCTG ATGAAACTGT CTTAAGAGTT TATGGCTTGT CTTTCTGAT | 1140 |
| 40 | TCTAGCTAGT AAAGAGTAGA TAAATATGAA ATATGTTTC CCTTTCTAG TTATGGTCAT | 1200 |
| | GGTTGGCTGG TATTCAATTTC TTTTATGGCA ATACTTGCTT CTAACTATCT TTAGTAGATT | 1260 |
| | CATGTATTAA CTTGTGAGTC ATTACTTTAT GGGTGTAGGG ATGCACAAACT TATTGGTGAC | 1320 |
| 45 | TTCAACAACT GGAATGGCTC TGGGCACAGG ATGACAAAGG ATAATTATGG TGTTGGTCA | 1380 |
| | ATCAGGATTT CCCATGTCAA TGGGAAACCT GCCATCCCC ATAATTCCAA GGTTAAATT | 1440 |
| 50 | CGATTTCAAC GTGGAGATGG ACTATGGTC GATCGGGTTC CTGCATGGAT TCGTTATGCA | 1500 |
| | ACTTTGATG CCTCTAAATT TGGAGCTCCA TATGACGGTG TTCACTGGGA TCCACCTTCT | 1560 |
| | GGTGAAAGGT CTACTTTAG TGGCTCGAGA GCAAGAAATC TAAGTAAAC CCACACAATT | 1620 |
| 55 | AACTTACATT AATGTGGAGA CATGATACTT TTATTGCTCG TTTTGCAGGT ATGTGTTAA | 1680 |
| | GCATCCTCGG CCTCGAAAGC CTGACGCTCC ACGTATTAC GAGGCTCATG TGGGGATGAG | 1740 |
| 60 | TGGTAAAAG CCTGAAGTAA GCACATACAG AGAATTGCA GACAATGTGT TACCGCCAT | 1800 |
| | AAAGGCAAAC AACTACAACA CAGTCAGCT GATGGCAATC ATGGAACATT CATATTATGC | 1860 |
| | TTCTTTGGG TACCATGTGA CGAATTCTT CGCAGTTAGC AGCAGATCAG AACGCCAGAG | 1920 |
| 65 | ACCTCAATAT CTTGTTGACA AGGCACATAG TTTACGGTTG CGTGTCTGA TGGATGTTGT | 1980 |
| | CCATAGCCAT GCGAGCACTA ATAAGACAGA TGGTCTTAAT GGCTATGATG TTGGGCAAAA | 2040 |

- 73 -

CACACAGGAG TCCTATTTCC ACACAGGAGA AAGGGGCTAT CATAAACTGT GGGATAGCCG 2100
CCTGTTCAAC TATGCCAATT GGGAGTCTTA CGATTTCTTC TTTCTAATCT GAGATATTGG 2160
5 ATGGACGAAT TCATGTTGA TGGCTTCCGA TTTGATGGGG TAACATCCAT GCTATATAAT 2220
CACCATGGTA TCAATATGTC ATTGCGCTGGA AGTTACAAGG AATATTTGG TTTGGATACT 2280
10 GATGTAGATG CAGTTGTTA CCTGATGCTT GCGAACCAATT TAATGCACAA ACTCTTGCCA 2340
GAAGCAACTG TTGTTGCAGA AGATGTTCA GGCATGCCAG TGCTTGTGCG GTCAGTTGAT 2400
15 GAAGGTGGAG TAGGGTTGA CTATGCCCTG GCTATGGCTA TTCTGATAG ATGGATCGAC 2460
TACTTGAAGA ACAAAAGATGA CCTTGAATGG TCAATGAGTG GAATAGCACA TACTCTGACC 2520
AACAGGAGAT ATACGGAAAA GTGCATTGCA TATGCTGAGA GCCATGATCA CGTATGTTT 2580
20 CCCTCCCTTG TCGCTGTGCG TGAGTATGTG TTCTTTTTT ATGGGGCACT GGTCTAAGAA 2640
CATACAGTTC AAAGGTGAGA CACTTTCTTT GCCTGGTAGA CAAATTGAG AAATAAACAT 2700
TTCGCTTGAT GACTTTAGT TGCTTCACAA GTTCGAATTA AGTTAGTTAT ATTCTGATAA 2760
25 CTAGTGATAG TACCCACTAA CCAGCTATTA CGGACCATGT AAGAATGTCC GAAGACTGCA 2820
GTTATATATC GTTGACTTTG TGTCATCTA TTGAAACAAC TTAGTAGTTA ACTTTCACGC 2880
30 AAATTTCAAG TCTATTGTTG GCGACAAGAC TATGGCATTCTT CTCTTGATGG ACAAGGAAAT 2940
GTATACTGGC ATGTCAGACT TGCAGCCTGC TTGCGCTACA ATTGATCGTG GAATTGCACT 3000
35 TCAAAAGGTT CGATTCGTTT TAAGTATTCC TGAATTGAT GTTCTAGTTC CAGACGAGTA 3060
TTGTAATGTT CGTTGTTACT CAGAGTTCTG CTTAGTCCTT GAAGATAATG TATTCCAGTC 3120
CCTTTGGTA CATTGGCTT ATTTGTTAC AAATATTCA GATGATTCAC TTCATCACCA 3180
40 TGGCCCTTGG AGGTGATGGC TACTTGAATT TTATGGTAA TGAGGTAATA TCTGGTTATC 3240
TGTCAAAACT TATTCTGAT CAATATGTTT CGGGATTCCC TCGAAAAAAA TCCTTGGGC 3300
45 AGGGCGAAAA GTTTAACAT CTGTTTCTA TGATAGCCAA GTACTCCCCA GCTATTCCA 3360
TGTTATCACG TATCATTAG CTGTGCCGGT AGTTAATCTT TATTCTAATT CATTGGTT 3420
TTTTAGCGTG GCAGTCTATT GTTGGATCCT CTTATTCCAA TTACATATAT GCCGACATCA 3480
50 CACACTTATG AATATTCCCT GTTTAAAAGA TTTTATTTT ATACCAATGT TTCTCCGTAA 3540
ATGATGCAAAC CATGATAGAG ATGTTAGCAT GTCTTCTTA ACCTACTCAT GTTTACATA 3600
TCACGACAAG CTTCTTGCAG AAAATCAGCA GTATATGGCA AATTGCTGCA ACCTGACAAC 3660
55 GTTTATATCT GTTTCTAAC TCATACTGAC GGTGCAATT CTTTTAGTT TGGCCACCCA 3720
GAATGGATTG ACTTTCCAGA AGAAGGCAAC AACTGGAGTT ATGATAAAATG CAGACGCCAG 3780
60 TGGAGCCTCG CAGACATTGA TCACCTACGA TACAAGGTTA TGCCTATGTA TATTGTTACA 3840
GTTTCTGGTC TGGTAGCTCT CTTGGGATCT TGACCTCACT TAGTTCCCTC ATCTCTGACT 3900
65 GTAGCTTATT TACACTGTGT TCCAACCTCT GTCTTGTGGA TAAATTCTCC CTTCTAACGT 3960
TTCATATTAA GCCTTCAAA CTAAACTAAA TTGCTGATCT ACTACTAGTT GCTCAGTACG 4020

- 74 -

ATGCCCCAAT CTTGCCTGTG GTCACCTAGT AATTTCTTG ATTCTTACAC ATTAGTGATA 4080
 TGCAAGTGCAT ACATTATCCA TATAAATTGA CATTGCAATT TCCCAAATAT TATTGAGG 4140
 5 CTGTGTTCTT TTGTTAACAG GAAGTTATTT TCTCTGCATC TGATAAAATAA TAATAGCCTT 4200
 TCACCGATTTC TCTCATATTT TATCCAACCTT TTCTGCATTC AAGCATTTC TGTTCCTCGC 4260
 10 CTACATATA TAATTTGAAC AGTACATGAA CGCATTGAT CAAGCAATGA ATGCGCTCGA 4320
 CGACAAAATTT TCCTTCCTAT CATCATCAAA GCAGATTGTC AGCGACATGA ATGAGGAAAA 4380
 GAAGTAGTTA ACTATACAAT GTTGTAGTCAG GGCAGCTGTT GCATCATTG ATTCACTCCT 4440
 15 ACTCTTAAGA ATAGCAACTC TGACTGTGC GTTTATGTT ACCAAATAAG TTGAAACCGT 4500
 ATCTGTTGA TATGAACCAT TGTTGTCTCA AAATGGGCTA TGGACTCAAT CCAACTTCCT 4560
 20 TTCCAGATTA TTGTATTGAA ACCTGGAATC TGGTCTTCGT CTTCAATTTC CATCCCAGTA 4620
 AAACTTATGA TGGGTAACGT ATCTCTTGCA AGCTTGCCT TTCAATATT CTTCTGCTTA 4680
 ATGACTAATG TGCTTAATCT CGTTTCCACT TTTAAAACAC GCAGTTACAA AGTCGGATGT 4740
 25 GACTTGCCCTG GGAAGTACAA GGTAGCTCTG GACTCTGATG CTCTGATGTT TGGTGGACAT 4800
 GGAAGAGTAA GCAATGTTAA TGATGTTCAA GATCTTTTT GCAACACTAT GTTCTTCTAT 4860
 30 AGAAGGGGCC ATCAAGGCTG CATCAGATAA TCTTATTGTC AGTGTGATC TGTGCTGCAT 4920
 CGCAGGTGGC CCATGACAAC GATCACTTTA CGTCACCTGA AGGAGTACCA GGAGTACCTG 4980
 AAACAAACTT CAACAACCGC CCTAACTCAT TCAAAATCCT GTCTCCATCC CGCACTTGTG 5040
 35 TGGTAATGCT AATTACTAGG AGGATTTAGT AACAAATAAT AAATAACAGC AAAAGATATC 5100
 TGCAGTACGA TCTCACAAAA TGCTCTCTTG CCAGGCTTAC TATCGCGTGC AGGAGAAAGC 5160
 40 GGAAGAGCCC AAGGATGAAG GAGCTGCTTT CTTGGGGAA ACTGCTCTCG GGTACATCGA 5220
 TGTTGAAGGCC ACTGGCGTCA AAGACGCAGC AGATGGTGAG GCGACTTCTG GTTCCGAAAA 5280
 GGCCTCTACA GGAGGTGACT CCAGCAAGAA GGGATTAAAC TTTGTCTTC TGTCACCCGA 5340
 45 CAAAGACAAC AAATAAGCAC CATATCAACG CTTGATCAGG ACCGTGTGCC GACGTCCCTG 5400
 TAATACTCCT GCTATTGCTA GTAGTAGCAA TACTGTCAA CTGTGCAGAC TTGAAATTCT 5460
 50 GGCTGGACT TTGCTGAGGT TACCTACTAT ATAGAAAGAT AAATAAGCGG TGATGGTGCG 5520
 GGTGAGTCC AGCTATATGT GCCAAATATG CGCCATCCCG AGTCCTCTGT CATAAAAGAAA 5580
 GTTTCGGCT TCCATCCCAG AATAAAAACA GTTGTCTGTT TGCAATTCT TTTTGTCTTG 5640
 55 CATAGTTACA TGATAATTGA TGCAATTGTC TATAAGCCTG GATTGCATCT TCTTTGCTA 5700
 ATAACTGCAG GCCCAAGAAA GCCTAGATTG TATCTTTTT TGCTAATAAC TGCAGTGCTG 5760
 60 GGGAGCTTC AGTCCTTGT TCCGTTCTCG AGACAAGCGC TCATGTTGG CGCACAAAGG 5820
 TAAGCCATCA TCTTATCAAG TCCAAAATT CTCTGGTTGA AAGAAACCAT CACTAACTTG 5880
 TTCCAGGTGT TGGTTCTCC ACAACCAAAA GGCGACCATC GTCGTCATCA TCGCTCACAG 5940
 65 CACTGACCAT CGAAGCCACG GTGGGCATGA AATGCGCATC GCCCAAGACT TGGGACCGTT 6000
 TCAAAATATC ACAAAACTGCC ATGGCATCTT CTGCCAAAGG CTGCACTGCA CCTTTGGCAT 6060

- 75 -

| | | | |
|----|--------------------------------------------------------------------------------------------|------|-----|
| | GAACAGAAGC AACAGGGGCT TGGAACTGAA CGCCGAAAAT AAAGTCAAAC CGGCTGGGCC | 6120 | |
| 5 | GGATTGAAAG GGGAAACGCC AAAATCCACT TAATTTGAAT GGAAGGAGGA ATGGTTCTTG | 6180 | |
| | CTGGTTCAA CTCTGCAGGC TTCCCTCTGA ATTCACACG GAGCCATT | 6228 | |
| | (2) INFORMATION FOR SEQ ID NO: 10: | | |
| | (i) SEQUENCE CHARACTERISTICS: | | |
| 10 | (A) LENGTH: 11463 base pairs | | |
| | (B) TYPE: nucleic acid | | |
| | (C) STRANDEDNESS: single | | |
| | (D) TOPOLOGY: linear | | |
| 15 | (ii) MOLECULE TYPE: cDNA | | |
| | (iii) HYPOTHETICAL: NO | | |
| 20 | (iv) ANTI-SENSE: | | |
| | (vi) ORIGINAL SOURCE: | | |
| | (A) ORGANISM: tritum tauschii | | |
| | (F) TISSUE TYPE: Endosperm | | |
| 25 | (ix) FEATURE: | | |
| | (A) NAME/KEY: misc_feature | | |
| | (B) LOCATION: 1..11463 | | |
| | (D) OTHER INFORMATION:/product= "complete sequence of the starch branching enzyme II gene" | | |
| 30 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10: | | |
| | AGAAACACCT CCATTTAGA TTTTTTTTTT GTTCTTTTCG GACGGTGGGT CGTGGAGAGA | | 60 |
| 35 | TTAGCGTCTA GTTTCTTAA AAGAACAGGC CATTAGGCC CTGCTTACA AAAGGCTCAA | | 120 |
| | CCAGTCCAAA ACGTCTGCTA GGATCACCAAG CTGCAAAGTT AAGCGCGAGA CCACCAAAAC | | 180 |
| 40 | AGGCGCATTG GAACTGGACA GACGCTCACG CAGGAGCCCA GCACCCACAGG CTTGAGCCTG | | 240 |
| | ACAGCGGACG TGAGTGCCTG ACACATGGGG TCATCTATGG GCGTCGGAGC AAGGAAGAGA | | 300 |
| | GACGCACATG AACACCATGA TGATGCTATC AGGCCTGATG GAGGGAGCAA CCATGCACCT | | 360 |
| 45 | TTTCCCCCTCT GGAAATTCTAGT AGCTCACACT TTTTTTTAAT GGAAGCAAGA GTTGGCAAAC | | 420 |
| | ACATGCATTT TCAAACAAGG AAAATTAATT CTCAAACCAC CATGACATGC AATTCTCAA | | 480 |
| 50 | CCATGCACCG ACCAGTCCAT GCGAGGTGGA AACGAAGAAC TGAAAATCAA CATCCCAGTT | | 540 |
| | GTCGAGTCGA GAAGAGGATG AACTGAAAG TATGCGTATT ACGATTCAT TTACATACAT | | 600 |
| | GTACAAATAC ATAATGTACC CTACAATTG TTTTTGGAG CAGAGTGGTG TGGTCTTTT | | 660 |
| 55 | TTTTTACACG AAAATGCCAT AGCTGGCCCG CATGCGTGCA GATCGGATGA TCGGTGGAG | | 720 |
| | ACGACGGACA ATCAGACACT CACCAACTGC TTTTGTCTGG GACACAATAA ATGTTTTGT | | 780 |
| 60 | AAACAAAATA AATACTTATA AACGAGGGTA CTAGAGGCCG CTAACGGCAT GGCCAGGTAA | | 840 |
| | ACGCGCTCCC AGCCGTTGGT TTGCGATCTC GTCCTCCCGC ACGCAGCGTC GCCTCCACCG | | 900 |

- 76 -

TCCGTCCGTC GCTGCCACCT CTGCTGTGCG CGCGCACGA. GGGAGGAAGA ACGAACGCCG 960
 CACACACACT CACACACGGC ACACCCCCCG TGGTCCCTT TTCCGGCTTG GCCTATCT 1020
 5 CCTCTCCCCC GCCCATCCCC ATGCACTGCA CCGTACCCGC CAGCTTCCAC CCCGCCGCA 1080
 CACGTTGCTC CCCCTCTCA TCGCTTCTCA ATTAATATCT CCATCACTCG GGTCGGCGC 1140
 10 TGCATTTCGG CGGGCGGGTT GAGTGAGATC TGGCGACTG GCTGACTCAA TCAC TACGCG 1200
 GGGATGGCGA CGTCGCGGT GTCCGGCGCG ACTCTCGGTG TGGCGCGGC CGGCCTCGGA 1260
 GTGGCGCGGG CGGGCTCGGA CGGGAGGGC GGGCGGGACT TGCCGTCGCT GCTCCTCAGG 1320
 15 AAGAAGGACT CCTCTCGTAC GCCTCGCTCT CTCGAATCTC CCCCCTCTGG CTTGGCTCC 1380
 CCTTCTCTCT CCTCTCGCG CGCATGGCCT GTTCGATGCT GTTCCCCAAT TGATCTCCAT 1440
 GAGTGAGAGA GATAGCTGGA TTAGGCGATC GCGCTCCTG AACCTGTATT TTTTCCCCCG 1500
 20 CGGGGAAATG CGTTAGTGTC ACCCAGGCC TGGTGT TACC ACGGCTTGA TCATTCCTCG 1560
 TTCATTCTG ATATATATTT TCTCATTCTT TTTCTTCCTG TTCTTGCTGT AACTGCAAGT 1620
 25 TGTGGCGTTT TTTCACTATT GTAGTCATCC TTGCATTTG CAGGCGCCGT CCTGAGCCGC 1680
 GCGGCCTCTC CAGGGAAAGGT CCTGGTGCCT GACGGCGAGA GGACGACTTG GCAAGTCCGG 1740
 30 CGCAACCTGA AGAATTACAG GTACACACAC TCGTCCCGT AAATCTTCAT ACAATCGTTA 1800
 TTCACCTTACC AAATGCCGGA TGAAACCAAC CACGGATGCG TCAGGTTTCG AGCTTCTTCT 1860
 ATCAGCATTG TGCAGTACTG CACTGCCTTG TTCATTGCT TAGCCTTGGC CCCGTGCTGG 1920
 35 CTCTGGGCC ACTGAAAAAA TCAGATGGAT GTGCATTCTA GCAAGAACTT CACAACATAA 1980
 TGCACCGTTT GGGGTTTCGT CAGTCTGCTC TACAATTGCT ATTTTCTGT CTGTAGATAC 2040
 40 CTGAAGATAT CGAGGAGCAA ACGGCGGAAG TGAACATGAC AGGGGGGACT GCAGAGAAAC 2100
 TTCAATCTTC AGAACCGACT CAGGGCATTG TGGAAACAAT CACTGATGGT GTAACCAAAG 2160
 GAGTTAAGGA ACTAGTCGTG GGGGAGAAC CGCGAGTTGT CCCAAAACCA GGAGATGGC 2220
 45 AGAAAATATA CGAGATTGAC CCAACACTGA AAGATTTCG GAGCCATCTT GACTACCGGT 2280
 AATGCCTACC CGCTGCTTTC GCTCATTG AATTAAGGTC CTTTCATCAT GCAAATTGG 2340
 50 GGAACATCAA AGAGACAAAG ACTAGGGACC ACCATTTCAT ACAGATCCCT TCGTGGTCTG 2400
 AGAATATGCT GGGAAAGTAA TGTATAATTG ATGGCTACAA TTTGCTAAA ATTGCAATAC 2460
 GAATAACTGT CTCCGATCAT TACAATTAAA GAGTGGCAA CTGATGAAAA TGTGGTGGAT 2520
 55 GGGTTATAGA TTTTACTTTG CTAATTCTC TACAAATTG CTAGGGGGGA AATCTACCAAG 2580
 TTGGGAAACT TAGTTCTTA TCTTGTGGC CTTTTGTTT TGGGGAAAAC ACATTGCTAA 2640
 60 ATTGAAATGA TTTTGGGTAT ACCTCGGTGG ATTCAACAGA TACAGCGAAT ACAAGAGAAT 2700
 TCGTGCTGCT ATTGACCAAC ATGAAGGTGG ATTGGAAGCA TTTTCTCGTG GTTATGAAAA 2760
 GCTTGGATTG ACCCGCAGGT AAATTAAAG CTTTATTATT ATGAAACGCC TCCACTAGTC 2820
 65 TAATTGCATA TCTTATAAGA AAATTATAAA TTCCTGTTT CCCCTCTCTT TTTCCAGTG 2880
 CTGAAGGTAT CGTCTAATTG CATATCTTAT AAGAAAATTT ATATTCTGT TTTCCCTAT 2940

- 77 -

TTTCCAGTGC TGAAGGTATC ACTTACCGAG AATGGGCTCC CTGGAGCGCA TGTTATGTTC 3000
5 TTTTAAGTTC CTTAACGAGA CACCTTCCAA TTTATTGTTA ATGGTCACTA TTCACCAACT 3060
AGCTTACTGG ACTTACAAAT TAGCTTACTG AATACTGACC AGTTACTATA AATTATGAT 3120
CTGGCTTTG CACCCCTGTTA CAGTCTGCAG CATTAGTAGG TGACTTCAAC AATTGGAATC 3180
10 CAAATGCAGA TACTATGACC AGAGTATGTC TACAGCTTGG CAATTTCCA CCTTGCTTC 3240
ATAACTACTG ATACATCTAT TTGTATTTAT TTAGCTGTTT GCACATTCCCT TAAAGTTGAG 3300
15 CCTCAACTAC ATCATATCAA AATGGTATAA TTTGTCACTG TCTTAAGCTT CAGCCCAAAG 3360
ATTCTACTGA ATTTAGTCCA TCTTTTGAG ATTGAAAATG AGTATATTAA GGATGAATGA 3420
ATACGTGCAA CACTCCCATC TGCAATTATGT GTGCTTTCC ATCTACAATG AGCATATTTC 3480
20 CATGCTATCA GTGAAGGTTT GCTCCTATTG ATGCAGATAT TTGATATGGT CTTTCAGGA 3540
TGATTATGGT GTTTGGGAGA TTTTCCTCCC TAACAACGCT GATGGATCCT CAGCTATTCC 3600
25 TCATGGCTCA CGTGTAAAGG TAAGCTGGCC AATTATTTAG TCGAGGATGT AGCATTTCG 3660
AACTCTGCCT ACTAAGGGTC CCTTTTCCTC TCTGTTTTT AGATACGGAT GGATACTCCA 3720
TCCGGTGTGA AGGATTCAAT TTCTGCTTGG ATCAAGTTCT CTGTGCAGGC TCCAGGTGAA 3780
30 ATACCTTCA ATGGCATATA TTATGATCCA CCTGAAGAGG TAAGTATCGA TCTACATTAC 3840
ATTATTAAAT GAAATTCCA GTGTTACAGT TTTTAATAC CCACCTCTTA CTGACATGTG 3900
35 AGTCAAGACA ATACTTTGA ATTTGGAAGT GACATATGCA TTAATTCAACC TTCTAAGGGC 3960
TAAGGGGCAA CCAACCTTGG TGATGTGTGT ATGCTTGTGT GTGACATAAG ATCTTATAGC 4020
TCTTTATGT GTTCTCTGTT GGTTAGGATA TTCCATTTG GCCTTTGTG ACCATTACT 4080
40 AAGGATATT ACATGCAAAT GCAGGAGAAG TATGCTTCC AACATCTCAA CAAACGACC 4140
AGAGTCACTA AGGATTTATG AATCACACAT TGGAATGAGC AGCCCGGTAT GTCAATAAGT 4200
TATTCACCT GTTTCTGGTC TGATGGTTA TTCTATGGAT TTTCTAGTTC TGTTATGTAC 4260
45 TGTTAACATA TTACATGGTG CATTCACTTG ACAACCTCGA TTTTATTTTC TAATGTCTTC 4320
ATATTGGCAA GTGAAAAGT TTGCTTCCCTC TTTGTCTGCT TGTTCTTTG TCTTCTGTAA 4380
50 GATTTCCATT GCATTGGAG GCAGTGGCA TGTGAAAGTC ATATCTATT TTTTTTGTC 4440
AGAGCATACT TATATGAATT CCATTGTTGT TGCAATAGCT CGGTATAATG TAACCATGTT 4500
55 ACTAGCTTAA GATTTCCCAC TTAGGATGTA AGAAATATTG CATTGGAGCG TCTCCAGCAA 4560
GCCATTCCT ACCTTATTAA TGAGAGAGAG ACAAGGGGGGG GGGGGGGGGGG GGGGTTCCCT 4620
TCATTATTCT GCGAGCGATT CAAAAACTTC CATTGTTCTG AGGTGTACGT ACTGCAGGGA 4680
60 TCTCCCATTA TGAAGAGGAT ATAGTTAATT CTTTGTAAACC TACTTGGAAA CTTGAGTCTT 4740
GAGGCATCGC TAATATATAC TATCATCACA ATACTTAGAG GATGCATCTG AAATTTAGT 4800
65 GTGATCTTGC ACAGGAACCG AAGATAAATT CATATGCTAA TTTTAGGGAT GAGGTGTTC 4860
CAAGAATTAA AAGGCTTGGG TACAATGCAG TGCAGATAAT GGCAATCCAG GAGCATTCA 4920

- 78 -

ACTATGCAAG CTTGGGTAT TCACACAAATC CATTTCCTTC TGTATAACT CTTCACCCAT 4980
 TTGGAGCTAT TACATCCTAA TGCTTCATGC ACATAAAATA TTTGGATATA ATCCTTATT 5040
 5 AGATATATAG TACAACATACA CTTAGTATTC TGAAAAAGAT CATTATATTG TTGTTGGCTT 5100
 GTTCCAGGTA CCATGTTACT AATTTTTTG CACCAAGTAG CCGTTTGGA ACTCCAGAGG 5160
 10 ACTTAAAATC CTTGATCGAT AGAGCACATG AGCTTGGTT GCTTGTCTT ATGGATATTG 5220
 TTCATAGGTA ATTAGTCCAA TTTAATTAA GCTGTTTAC TGTTTATCTG GTATTCTAAA 5280
 GGGAAATTCA GGCAATTATG ATACATTGTC AAAAGCTAAG AGTGGCGAAA GTGAAATGTC 5340
 15 AAAATCTAGA GTGGCATAAG GAAAATTGGC AAAAACTAGA GTGGCAAAA TAAAATTTC 5400
 CCATCCTAAA TGGCAGGGCC CTATGCCGA ATATTTTCC ATTCTATATA ATTGTGCTAC 5460
 20 GTGACTTCCTT TTTTCTCAGA TGTATTAAAC CAGTTGGACA TGAAATGTAT TTGGTACATG 5520
 TAGTAAACTG ACAGTTCCAT AGAATATCGT TTTGTAATGG CAACACAATT TGATGCCATA 5580
 .. GATGTGGATT GAGAAGTTCA GATGCTATCA ATAGAATTAA TCAACTGGCC ATGTACTCGT 5640
 25 GGCACTACAT ATAGTTGCA AGTTGGAAAA CTGACAGCAA TACCTCACTG ATAAGTGGCC 5700
 AGGCCCACT TGCCAGCTTC ATACTAGATG TTACTTCCCT GTGAAATTCA TTTGAACATA 5760
 30 TTACTTAAAG TTCTTCATTT GTCCTAACGTC AAACTTCTTT AAGTTTGACC AAGTCTATTG 5820
 GAAAATATAT CAACATCTAC AACACAAAT TACTTTGATC AGATTAACAA TTTTATTAA 5880
 ATTATATTAG CACATCTTG ATGTTGTTAGA TATCAGCACA TTTTCTATA GACTTGGTCA 5940
 35 AATATAGAGA AGTTTGACTT AGGACAAATC TAGAACTTCA ATCAATTGG ATCAGAGGGA 6000
 ACATCAAATA ATATAGATAG ATGTCAACAC TTCAACAAAA AAATCAGACC TTGTCACCAT 6060
 40 ATATGCATCA GACCACATGT TTGCTTCTAGC CACTTGCTT CATATTATG TGTTTGTACC 6120
 TAATCTACTT TTCCCTCTAC TTGGTTGGT TGATTCTATT TCAGTTGCAT TGCTTCATCA 6180
 ATGATTTGT GTACCCCTGCA GTCATTCGTC AAATAATACC CTTGACGGTT TGAATGGTTT 6240
 45 CGATGGCACT GATACACATT ACTTCCACGG TGGTCCACGC GGCCATCATT GGATGTGGGA 6300
 .. TTCTCGTCTA TTCAACTATG GGAGTTGGGA AGTATGTAGC TCTGACTTCT GTCACCATAT 6360
 50 TTGGCTAACT GTTCCTGTAA ATCTGTTCTT ACACATGTTG ATATTCTATT CTTATGCAGG 6420
 TATTGAGATT CTTACTGTCA AACCGGAGAT GGTGGCTTGA AGAATATAAG TTTGATGGAT 6480
 TTCGATTTGA TGGGGTGACC TCCATGATGT ATACTCACCA TGGATTACAA GTAAGTCATC 6540
 55 AAGTGGTTTC AGTAACCTTT TTAGGGCACT GAAACAATTG CTATGCATCA TAACATGTAT 6600
 CATGATCAGG ACTTGTGCTA CGGAGTCTTA GATAGTTCCC TAGTATGCTT GTACAATTAA 6660
 60 ACCTGATGAG ATCATGGAAG ATTGGAAGTG ATTATTATT ATTTCCTTTC TAAGTTGTT 6720
 TCTTGTCTA GATGACATT ACTGGGAACG ATGGCGAATA TTTTGGATTT GCTACTGATG 6780
 TTGATGCGGT AGTTTACTTG ATGCTGGTCA ACGATCTAAT TCATGGACTT TATCCTGATG 6840
 65 CTGTATCCAT TGGTGAAGAT GTAAGTGCTT ACAGTATTAA TGATTTTAA CTAGTTAAGT 6900
 AGTTTATTAA TGGGGATCAG TCTGTTACAC TTTTGTAG GGGTAAATC TCTCTTTCA 6960

TAACAATGCT AATTATACCC TTGTATGATA ATGCATCACT TAGTAATTTG AAAAGTGCAA 7020
GGGCATTCAA GCTTACGAGC ATATTTTTG ATGGCTGTAA TTTATTTGAT AGTATGCTTG 7080
5 TTTGGGTTTT TCAATAAGTG GGAGTGTGTG ACTAATGTTG TATTATTTAT TTAATTGCGG 7140
AAGAAATGGG CAACCTTGTC AATTGCTTCA GAAGGCTAAC TTTGATTCCA TAAACGCTTT 7200
10 GGAAATGAGA GGCTATTCCC AAGGACATGA ATTATACCTC AGTGTGTTCT GTACATGTAT 7260
TTGTAATAGT GGTTTAACCT AAATTCCCTGC ACTGCTATGG AATCTCACTG TATGTTGTAG 7320
15 TGTACACATC CACAAACAAG TAATCCTGAG CTTTCAACTC ATGAGAAAAT AGAGTCCGCT 7380
TCTGCCAGCA TTAACTGTTAC ACAGTTCTAA TTTGTGTAAC TGTGAAATTG TTCAGGTCAG 7440
TGGAATGCCT ACATTTGCA TCCCTGTTCC AGATGGTGGT GTTGGTTTTG ACTACCGCCT 7500
20 GCATATGGCT GTAGCAGATA AATGGATTGA ACTCCTCAAG TAAGTGCAGG AATATTGGTG 7560
ATTACATGCG CACAATGATC TAGATTACAT TTTCTAAATG GTAAAAAGGA AAATATGTAT 7620
25 GTGAATATCT AGACATTTCCTGC CTGTTATCAG CTTGAATAACG AGAAGTCAAA TACATGATTT 7680
AAATAGCAAA TCTCGGAAAT GTAATGGCTA GTGTCTTTAT GCTGGCAGT GTACATTGCG 7740
CTGTAGCAGG CCAGTCAAACA CAGTTAGCAA TATTTTCAGA AACAAATATTA TTTATATCCG 7800
30 TATATGAGAA AGTTAGTATA TAAACTGTGG TCATTAATTG TGTTCACCTT TTGTCCTGTT 7860
TAAGGATGGG CAGTAGGTAA TAAATTAGC CAGATAAAAAT AAATCGTTAT TAGTTTACA 7920
35 AAAGGAATAT ACAGGGTCAT GTAGCATATC TAGTTGTAAT TAATGAAAAG GCTGACAAAA 7980
GGCTCGGTAA AAAAAGCTTT ATGATGATCC AGATAGATAT GCAGGAACGC GACTAAAGCT 8040
CAAATACTTA TTGCTACTAC ACAGCTGCCA ATCTGTACATG ATCTGTGTT TGCTTTGTGC 8100
40 TATTTAGATT TAAATACTAA CTCGATACAT TGGCAATAAT AAACCTTAAC ATTCAACCAA 8160
TTTGGTGGAT ACCAGAATTCTGCTCTT GTTAGTAATG ATGTGCTCCC TGCTGCTGTT 8220
CTCTGCCGTT ACAAAAGCTG TTTTCAGTTT TTTGCATCAT TATTTTTGTG TGTGAGTAGT 8280
45 TTAAGCATGT TTTTGAAAGC TGTGAGCTGT TGGTACTTAA TACATTCTTG GAAGTGTCCA 8340
AATATGCTGC AGTGTAAATT AGCATTCTT TAACACAGGC AAAGTGACGA ATCTTGGAAA 8400
50 ATGGCGATA TTGTGCACAC CCTAACAAAT AGAAGGTGGC TTGAGAAGTG TGTAACCTTAT 8460
GCAGAAAGTC ATGATCAAGC ACTAGTTGGT GACAAGACTA TTGCATTCTG GTTGATGGAT 8520
AAGGTACTAG CTGTTACTTT TGGACAAAAG AATTACTCCC TCCCGTTCCCT AAATATAAGT 8580
55 CTTTGAGAG ATTCCACTAT GGACCACATA GTATATAGAT GCATTTAGA GTGTAGATTC 8640
ACTCATTTCCTG CTTCGTATGT AGTCCATAGT GAAATCTCTA CAGAGACTTA TATTTAGGAA 8700
60 CGGAGGGAGT ACATAATTGA TTTGTCTCAT CAGATTGCTA GTGTTTCTT GTGATAAAGA 8760
TTGGCTGCCT CACCCATCAC CAGCTATTTC CCAACTGTTA CTTGAGCAGA ATTTGCTGAA 8820
AACGTACCAT GTGGTACTGT GGCGGCTTGT GAACTTTGAC AGTTATGTTG CAATTTCTG 8880
65 TTCTTATTTA TTTGATTGCT TATGTTACCG TTCATTTGCT CATTCCCTTC CGAGACCAGC 8940

- 80 -

CAAAGTCACG TGTTAGCTGT GTGATCTGTT ATCTGAATCT TGAGCAAATT TTATTAATAG 9000
GCTAAAATCC AACGAATTAT TTGCTTGAAT TTAAATATAC AGACGTATACTC ACACCTGGCT 9060
5 CTTTCTTAGA TGATTACCAT AGTGCCTGAA GGCTGAAATA GTTTTGGTGT TTCTGGATG 9120
CCGCCTAAAG GAGTGATTT TATTGGATAG ATTCCCTGGCC GAGTCTTCGT TACAACATAA 9180
10 CATTGGAG ATATGCTTAG TAACAGCTCT GGGAAAGTTG GTCACAAGTC TGCATCTACA 9240
CGCTCCTTGA GGTTTATTAG TGGGCCATC TTTGTAACTA GTGGCACCTG TAAGGAAACA 9300
CATTCAAAAG GAAACGGTCA CATCATTCTA ATCAGGACCA CCATACTAAG AGCAAGATTG 9360
15 TGTTCCAATT TTATGAGTTT TTGGGACTCC AAAGGAAACA AAAGTGTCTC ATATTGTGCT 9420
TATAACTACA GTTGTTTTA TACCAGTGTA GTTTTATTCC AGGACAGTTG ATACTTGGTA 9480
20 CTGTGCTGTA AATTATTTAT CCGACATAGA ACAGCATGAA CATATCAAGC TCTCTTGTG 9540
CAGGATATGT ATGATTTCAT GGCTCTGGAT AGGCTTCAAC TCTTCGCATT GATCGTGGCA 9600
TAGCATTACA TAAAATGATC AGGCTTGTCA CCATGGTTT AGGTGGTGAAGGCTATCTTA 9660
25 ACTTCATGGG AAATGAGTTT GGGCATCCTG GTCAGTCTTT ACAACATTAT TGCATTCTGC 9720
ATGATTGTGA TTTACTGTAA TTTGAACCAT GCTTTCTTT CACATTGTAT GTATTATGTAA 9780
30 ATCTGTTGCT TCCAAGGAGG AAGTTAACCTT CTATTTACTT GGCAGAAATGG ATAGATTTTC 9840
CAAGAGGCCAC AAAACTCTT CCAACCGGCA AAGTTCTCCC CTGGAAATAA CAATAGTTAT 9900
GATAAAATGCC GCCGTAGATT TGATCTTGTA AGTTTAGCT GTGCTATTAC ATTCCCTCAC 9960
35 TAGATCTTTA TTGGCCATT ATTCTTGAT GAAATCATAA TGTTTGTAG GAAAGATCAA 10020
CATTGCTTT GTAGTTTGT AGACGTTAAC ATAAGTATGT GTTGAGAGTT GTTGATCATT 10080
40 AAAAATATCA TGATTTTTG CAGGGAGATG CAGATTCTCT TAGATATCGT GGTATGCAAG 10140
AGTTCGATCA GGCAATGCAG CATCTTGAGG AAAAATATGG GGTATGTCAC TTGTTGTCT 10200
TTGTTGCATA ACAAGTCACA GTTTAACGTC AGTCTCTCA AGTGGTAAAA AAAGTGTAGA 10260
45 ATTAATTCT GTAATGAGAT GAAAATGTG CAAAGGCCGA GCTGGAATTG CTTTCACCA 10320
AAACTATTT CTTAAGTGT TGTGTATTGA TACATATACC AGCACTGACA ATGTAACCTGC 10380
50 AGTTTATGAC ATCTGAGCAC CAGTATGTTT CACGGAAACA TGAGGAAGAT AAGGTGATCA 10440
TCCTCAAAAG AGGAGATTG GTATTTGTTT TCAACTTCCA CTGGAGCAAT AGCTTTTTG 10500
ACTACCGTGT TGGGTGTTCC AAGCCTGGGA AGTACAAGGT ATGCTTGCCT TTTCATTGTC 10560
55 CACCCCTCAC CAGTAGGGTT AGTGGGGCT TCTACAACCTT TTAATTCCAC ATGGATAGAG 10620
TTTGTGGTC GTGCAGCTAT CAATATAAG AATAGGGTAA TTTGTAAAGA AAAGAATTG 10680
60 CTCGAGCTGT TGTAGCCATA GGAAGGTTGT TCTTAACAGC CCCGAAGCAGC ATACCATTCA 10740
TTCATATTAT CTACTTAAGT GTTTGTTCA ATCTTTATGC TCAGTTGGAC TCGGTCTAAT 10800
ACTAGAACTA TTTTCCGAAT CTACCCCTAAC CATCCTAGCA GTTTTAGAGC AGCCCCATTT 10860
65 GGACAATTGG CTGGGTTTTT GTTAGTTGTG ACAGTTCTG CTATTTCTTA ATCAGGTGGC 10920
CTTGGACTCT GACGATGCAC TCTTGGTGG ATTCAAGCAGG CTTGATCATG ATGTCGACTA 10980

- 81 -

CTTCACAACC GTAAGTCTGG GCTCAAGCGT CACTTGACTC GTCTTGACTC AACTGTTAC 11040
AAATCTGAAT CAACTTCCA ATTGCTGATG CCCTTGCAGG AACATCCGCA TGACAAACAGG 11100
5 CCGCGCTCTT TCTCGGTGTA CACTCCGAGC AGAACTGCGG TCGTGTATGC CCTTACAGAG 11160
TAAGAACCAAG CAGCGGCTTG TTACAAGGCA AAGAGAGAAC TCCAGAGAGC TCGTGGATCG 11220
TGAGCGAAGC GACGGGCAAC GGCGCGAGGC TGCTCCAAGC GCCATGACTG GGAGGGGATC 11280
10 GTGCCTCTTC CCCAGATGCC AGGAGGAGCA GATGGATAGG TAGCTTGTG 11340
GAAAGAAAAT GGACGGGCCT GGGTGTGTTGT TGTGCTGCAC TGAACCCTCC TCCTATCTG 11400
15 CACATTCCCG GTTGTGTTTG TACATATAAC TAATAATTGC CCGTGCCTC AACGTGAAAA 11460
TCC 11463

(2) INFORMATION FOR SEQ ID NO: 11:
20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2662 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
25 (ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
30 (iv) ANTI-SENSE:
(vi) ORIGINAL SOURCE:
(A) ORGANISM: triticum tauschii
(F) TISSUE TYPE: Endosperm
35 (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..2651
(D) OTHER INFORMATION:/product= "nucleotide sequence of
40 cDNA wheat SSS I"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

45 TCTCCCACTC TTCTCTCCCC GCGCACACCG AGTCGGCACC GGCTCATCAC CCATCACCTC 60
GGCCTCGGCC ACCGGAAAC CCCCCGATCC GCTTTGCAG GCAGCGCACT AAAACCCGG 120
GGAGCGCGCC CGCGGGCAGC AGCAGCACCG CAGTGGGAGA GAGAGGCTTC GCCCCGGCCC 180
50 GCACCGAGCG GGGCGATCCA CCGTCCGTGC GTCCGCACCT CCTCCGCCTC CTCCCCGTGC 240
CCGCGCGCCC ACACCCATGG CGGGCACGGG CGTCGGCGCC GGGTGCCTCG CCCCCAGCGT 300
55 CCGCCTGCGC GCCGATCCGG CGACGGCGGC CCGGGCGTCC GCCTGCGTGC TCCGCGCGCG 360
GCTCCGGCGC TTGGCGCGGG GCCGCTACGT TGCCGAGCTC AGCAGGGAGG GCCCCCGCGC 420
GCGCCCCCGCG CAGCAGCAGC AACTGGCCCC GCCGCTCGTG CCAGGCTTCC TCGCGCCGCC 480
60 GCCGCCCCCGCG CCCGCCCAGT CGCCGGCCCC GACGCAGCCG CCCCTGCCGG ACGCCGGCGT 540
GGGGGAACTC GCGCCCGACC TCCTGCTCGA AGGGATTGCT GAGGATTCCA TCGACAGCAT 600

| | | | | | | | |
|----|------------|-------------|-------------|-------------|-------------|------------|------|
| | AATTGTGGCT | GCAAGTGAGC | AGGATTCTGA | GATCATGGAT | GCGAATGAGC | AACCTCAAGC | 650 |
| 5 | TAAAGTTACA | CGTAGCATCG | TGTTTGTGAC | TGGTGAAGCT | GCTCCTTATG | CAAAGTCAGG | 720 |
| | GGGGCTGGGA | GATGTTGTG | GTTCGTTACC | AATTGCTCTT | GCTGCTCGTG | GTCACCGTGT | 780 |
| | GATGGTTGTA | ATGCCAAGAT | ACTTGAATGG | GTCCTCTGAT | AAAAACTATG | CAAAGGCATT | 840 |
| 10 | ATACACTGGG | AAGCACATTA | AGATTCCATG | CTTTGGGGA | TCACATGAAG | TGACCTTTT | 900 |
| | TCATGAGTAT | AGAGACAACG | TCGATTGGGT | GTTCGTCGAT | CATCCGTCA | ATCATAGACC | 960 |
| 15 | AGGAAGTTA | TATGGAGATA | ATTTGGTGC | TTTGTTGAT | AATCAGTTCA | GATACACACT | 1020 |
| | CCTTGCTAT | GCTGCATGCG | AGGCCCCACT | AATCCTTGAA | TTGGGAGGAT | ATATTATGG | 1080 |
| | ACAGAATTGC | ATGTTGTTG | TGAACGATTG | GCATGCCAGC | CTTGTGCCAG | TCCTCTTGC | 1140 |
| 20 | TGCAAAATAT | AGACCATAACG | GTGTTTACAG | AGATTCCCGC | AGCACCCCTG | TTATACATAA | 1200 |
| | TTTAGCACAT | CAGGGTCTGG | AGCCTGCAAG | TACATATCCT | GATCTGGGAT | TGCCACCTGA | 1260 |
| 25 | ATGGTATGGA | GCTTTAGAAT | GGGTATTTC | AGAATGGCA | AGGAGGCATG | CCCTTGACAA | 1320 |
| | GGGTGAGGCA | GTAACTTTT | TGAAAGGAGC | AGTCGTGACA | GCAGATCGAA | TTGTGACCGT | 1380 |
| | CAGTCAGGGT | TATTCATGGG | AGGTACAAAC | TGCTGAAGGT | GGACAGGGCC | TCAATGAGCT | 1440 |
| 30 | CTTAAGCTCC | CGAAAAAGTG | TATTGAATGG | AATTGTAAAT | GGAATTGACA | TTAATGATTG | 1500 |
| | GAACCCCACC | ACAGACAAGT | GTCTCCCTCA | TCATTATTCT | GTCGATGACC | TCTCTGGAAA | 1560 |
| 35 | GGCCAATGT | AAAGCTGAAT | TGCAGAAGGA | GCTGGTTTA | CCTGTAAGGG | AGGATGTTCC | 1620 |
| | TCTGATTGGC | TTTATTGGAA | GACTGGATTA | CCAGAAAGGC | ATTGATCTCA | TTAAAATGGC | 1680 |
| | CATTCCAGAG | CTCATGAGGG | AGGACGTGCA | GTTCGTCATG | CTTGGATCTG | GGGATCCAAT | 1740 |
| 40 | TTTGAAAGGC | TGGATGAGAT | CTACCGAGTC | GAGTTACAAG | GATAAATTCC | GTGGATGGGT | 1800 |
| | TGGATTTAGT | GTTCAGTTT | CCACACAGAAT | AACTGCAGGT | TGCGATATAT | TGTTAATGCC | 1860 |
| 45 | ATCCAGGTTT | GAACCTTGTG | GTCTTAATCA | GCTATATGCT | ATGCAATATG | GTACAGTTCC | 1920 |
| | TGTAGTTCAT | GGAACTGGGG | GCCTCCGAGA | CACAGTCGAG | ACCTTCAACCC | CTTTGGTGC | 1980 |
| | AAAAGGAGAG | GAGGGTACAG | GGTGGCGTT | CTCACCGCTA | ACCGTGGACA | AGATGTTGTG | 2040 |
| 50 | GGCATTGCGA | ACCGCGATGT | CGACATTCA | GGAGCACAAG | CCGTCCTGGG | AGGGGCTCAT | 2100 |
| | GAAGCGAGGC | ATGACGAAAG | ACCATACGTG | GGACCATGCC | GCCGAGCAGT | ACGAGCAGAT | 2160 |
| 55 | CTTCGAATGG | GCCTTCGTGG | ACCAACCCTA | CGTCATGTAG | ACGGGGACTG | GGGAGGTCGA | 2220 |
| | AGCGCGGGTC | TCCTTGAGCT | CTGAAGACAT | GTTCCCTCATC | CTTCCGCGGC | CCGGAAGGAT | 2280 |
| | ACCCCTGTAC | ATTGCGTTGT | CCTGCTACAG | TAGAGTCGCA | ATGCGCCTGC | TTGCTGGTC | 2340 |
| 60 | CGCCGGTTCG | AGAGTAGATG | ACGGCTGTGC | TGCTGCGCG | GTGACAGCTT | CGGGTGGATG | 2400 |
| | ACAGTTACAG | TTTGCGGGAA | TAAGGAAGGG | ATGTGCTGCA | GGATGGTTAA | CAGCAAAGCA | 2460 |
| 65 | CCACTCAGAT | GGCAGCCTCT | CTGTCGCGTGT | TACAGCTGAA | ATCAGAAACC | AACTGGTGAC | 2520 |
| | TCTTTAGCCT | TAGCGATTGT | GAAGTTGTT | GCATTCTGTG | TATGTTGTCT | TGTCCTTAGC | 2580 |

- 83 -

TGACAAATAT TAGACCTGTT GGAGAATTTC ATTTATCTTT GCTGCTGTTG TTTTTGTTTT 2640
GTTAAAAAAA AAAAAAAA AA 2662

5 (2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 768 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: protein

15 (iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: triticum tauschii

20 (ix) FEATURE:

- (A) NAME/KEY: Protein
- (B) LOCATION: 1..768

25 (ix) FEATURE:

- (A) NAME/KEY: Protein
- (B) LOCATION: 1..768
- (D) OTHER INFORMATION:/product= "deduced amino acid sequence SBE II"

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met Ala Thr Phe Ala Val Ser Gly Ala Thr Leu Gly Val Ala Arg Pro
1 5 10 15

35 Pro Ala Ala Ala Gln Pro Glu Glu Leu Gln Ile Pro Glu Asp Ile Glu
20 25 30

Glu Gln Thr Ala Glu Val Asn Met Thr Gly Gly Thr Ala Glu Lys Leu
35 40 45

40 Glu Ser Ser Glu Pro Thr Gln Gly Ile Val Glu Thr Ile Thr Asp Gly
50 55 60

Val Thr Lys Gly Val Lys Glu Leu Val Val Gly Glu Lys Pro Arg Val
65 70 75 80

45 Val Pro Lys Pro Gly Asp Gly Gln Lys Ile Tyr Glu Ile Asp Pro Thr
85 90 95

50 Leu Lys Asp Phe Arg Ser His Leu Asp Tyr Arg Tyr Ser Glu Tyr Arg
100 105 110

Arg Ile Arg Ala Ala Ile Asp Gln His Glu Gly Gly Leu Glu Ala Phe
115 120 125

55 Ser Arg Gly Tyr Glu Lys Leu Gly Phe Thr Arg Ser Ala Glu Gly Ile
130 135 140

60 Thr Tyr Arg Glu Trp Ala Pro Gly Ala His Ser Ala Ala Leu Val Gly
145 150 155 160

- 84 -

| | |
|----|-----------------------------------------------------------------|
| | Asp Phe Asn Asn Trp Asn Pro Asn Ala Asp Thr Met Thr Arg Asp Asp |
| | 165 170 175 |
| 5 | Tyr Gly Val Trp Glu Ile Phe Leu Pro Asn Asn Ala Asp Gly Ser Pro |
| | 180 185 190 |
| | Ala Ile Pro His Gly Ser Arg Val Lys Ile Arg Met Asp Thr Pro Ser |
| | 195 200 205 |
| 10 | Gly Val Lys Asp Ser Ile Ser Ala Trp Ile Lys Phe Ser Val Gln Ala |
| | 210 215 220 |
| 15 | Pro Gly Glu Ile Pro Phe Asn Gly Ile Tyr Tyr Asp Pro Pro Glu Glu |
| | 225 230 235 240 |
| | Glu Lys Tyr Val Phe Gln His Pro Gln Pro Lys Arg Pro Glu Ser Leu |
| | 245 250 255 |
| 20 | Arg Ile Tyr Glu Ser His Ile Gly Met Ser Ser Pro Glu Pro Lys Ile |
| | 260 265 270 |
| | Asn Ser Tyr Ala Asn Phe Arg Asp Glu Val Leu Pro Arg Ile Lys Arg |
| | 275 280 285 |
| 25 | Leu Gly Tyr Asn Ala Val Gln Ile Met Ala Ile Gln Glu His Ser Tyr |
| | 290 295 300 |
| 30 | Tyr Ala Ser Phe Gly Tyr His Val Thr Asn Phe Phe Ala Pro Ser Ser |
| | 305 310 315 320 |
| | Arg Phe Gly Thr Pro Glu Asp Leu Lys Ser Leu Ile Asp Arg Ala His |
| | 325 330 335 |
| 35 | Glu Leu Gly Leu Leu Val Leu Met Asp Ile Val His Ser His Ser Ser |
| | 340 345 350 |
| | Asn Asn Thr Leu Asp Gly Leu Asn Gly Phe Asp Gly Thr Asp Thr His |
| | 355 360 365 |
| 40 | Tyr Phe His Gly Gly Pro Arg Gly His His Trp Met Trp Asp Ser Arg |
| | 370 375 380 |
| | Leu Phe Asn Tyr Gly Ser Trp Glu Val Leu Arg Phe Leu Leu Ser Asn |
| 45 | 385 390 395 400 |
| | Ala Arg Trp Trp Leu Glu Glu Tyr Lys Phe Asp Gly Phe Arg Phe Asp |
| | 405 410 415 |
| 50 | Gly Val Thr Ser Met Met Tyr Thr His His Gly Leu Gln Met Thr Phe |
| | 420 425 430 |
| | Thr Gly Asn Tyr Gly Glu Tyr Phe Gly Phe Ala Thr Asp Val Asp Ala |
| | 435 440 445 |
| 55 | Val Val Tyr Leu Met Leu Val Asn Asp Leu Ile His Gly Leu His Pro |
| | 450 455 460 |
| | Asp Ala Val Ser Ile Gly Glu Asp Val Ser Gly Met Pro Thr Phe Cys |
| 60 | 465 470 475 480 |
| | Ile Pro Val Pro Asp Gly Gly Val Gly Phe Asp Tyr Arg Leu His Met |
| | 485 490 495 |

- 85 -

| | |
|----|-----------------------------------------------------------------|
| | Ala Val Ala Asp Lys Trp Ile Glu Leu Leu Lys Gln Ser Asp Glu Ser |
| | 500 505 510 |
| 5 | Trp Lys Met Gly Asp Ile Val His Thr Leu Thr Asn Arg Arg Trp Leu |
| | 515 520 525 |
| | Glu Lys Cys Val Thr Tyr Ala Glu Ser His Asp Gln Ala Leu Val Gly |
| | 530 535 540 |
| 10 | Asp Lys Thr Ile Ala Phe Trp Leu Met Asp Lys Asp Met Tyr Asp Phe |
| | 545 550 555 560 |
| | Met Ala Leu Asp Arg Pro Ser Thr Pro Arg Ile Asp Arg Gly Ile Ala |
| 15 | 565 570 575 |
| | Leu His Lys Met Ile Arg Leu Val Thr Met Gly Leu Gly Gly Glu Gly |
| | 580 585 590 |
| 20 | Tyr Leu Asn Phe Met Gly Asn Glu Phe Gly His Pro Glu Trp Ile Asp |
| | 595 600 605 |
| | Phe Pro Arg Gly Pro Gln Thr Leu Pro Thr Gly Lys Val Leu Pro Gly |
| | 610 615 620 |
| 25 | Asn Asn Asn Ser Tyr Asp Lys Cys Arg Arg Arg Phe Asp Leu Gly Asp |
| | 625 630 635 640 |
| | Ala Asp Phe Leu Arg Tyr His Gly Met Gln Glu Phe Asp Gln Ala Met |
| 30 | 645 650 655 |
| | Gln His Leu Glu Glu Lys Tyr Gly Phe Met Thr Ser Glu His Gln Tyr |
| | 660 665 670 |
| 35 | Val Ser Arg Lys His Glu Glu Asp Lys Val Ile Ile Phe Glu Arg Gly |
| | 675 680 685 |
| | Asp Leu Val Phe Val Phe Asn Phe His Trp Ser Asn Ser Phe Phe Asp |
| | 690 695 700 |
| 40 | Tyr Arg Val Gly Cys Ser Arg Pro Gly Lys Tyr Lys Val Ala Leu Asp |
| | 705 710 715 720 |
| | Ser Asp Asp Ala Leu Phe Gly Gly Phe Ser Arg Leu Asp His Asp Val |
| | 725 730 735 |
| 45 | Asp Tyr Phe Thr Thr Glu His Pro His Asp Asn Arg Pro Arg Ser Phe |
| | 740 745 750 |
| 50 | Ser Val Tyr Thr Pro Ser Arg Thr Ala Val Val Tyr Ala Leu Thr Glu |
| | 755 760 765 |

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10550 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

60 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

- (vi) ORIGINAL SOURCE:
(A) ORGANISM: triticum tauschii
- 5 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:1..316
(D) OTHER INFORMATION:/product= "exon 1"
- 10 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:1472..1828
(D) OTHER INFORMATION:/product= "exon 2"
- 15 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:2766..2823
(D) OTHER INFORMATION:/product= "exon 3"
- 20 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:2906..3028
(D) OTHER INFORMATION:/product= "exon 4"
- 25 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:4113..4194
(D) OTHER INFORMATION:/product= "exon 5"
- 30 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:4286..4459
(D) OTHER INFORMATION:/product= "exon 6"
- 35 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:4562..4643
(D) OTHER INFORMATION:/product= "exon 7"
- 40 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:4744..4855
(D) OTHER INFORMATION:/product= "exon 8"
- 45 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:4999..5021
(D) OTHER INFORMATION:/product= "exon 9"
- 50 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:5102..5192
(D) OTHER INFORMATION:/product= "exon 10"
- 55 (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION:8593..8718

(D) OTHER INFORMATION:/product= "exon 11"

(ix) FEATURE:

(A) NAME/KEY: exon

5 (B) LOCATION:8807..8915

(D) OTHER INFORMATION:/product= "exon 12"

(ix) FEATURE:

(A) NAME/KEY: exon

10 (B) LOCATION:8992..9104

(D) OTHER INFORMATION:/product= "exon 13"

(ix) FEATURE:

(A) NAME/KEY: exon

15 (B) LOCATION:9161..9199

(D) OTHER INFORMATION:/product= "exon 14"

(ix) FEATURE:

(A) NAME/KEY: exon

20 (B) LOCATION:9498..9713

(D) OTHER INFORMATION:/product= "exon 15"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

| | | |
|----|--------------------------------------------------------|-----|
| 25 | ATGGCGGCGA CGGGCGTCGG CGCCGGGTGC CTCGCCCA GCGTCCGCCT | 50 |
| | GCGGCCGAT CCGCGACGG CGGCCGGGC GTCCGCTTGC GTCGTCCGCG | 100 |
| 30 | CGCGGCTCCG GCGCTTGGCG CGGGGCCGCT ACGTCGCCGA GCTCAGCAGG | 150 |
| | GAGGGCCCCG CGGCGGCCCG CGCGCAGCAG CAGCAACTGG CCCCGCCGCT | 200 |
| | CGTGCCAGGC TTCCCTCGCGC CGCCGCCGCC CGCGCCCGCC CAGTCCGG | 250 |
| 35 | CCCCGACGCA GCCGCCCCCTG CCGGACGCCG GCGTGGGGGA ACTCGGCC | 300 |
| | GACCTCCTGC TCGAAGGTAA AAAACAAGGC TGAATCCTCA GATCACTCCG | 350 |
| | CGTCTCGTT TTACCAAATA CGGTACTGCG AAGTGGTGCT GTATATGTGA | 400 |
| 40 | AGTTTCTGTC GATTCTTCC TGACGGATGT TCAGTCGATT CAGTTGTATA | 450 |
| | TATGTGATAC GTTCGTTGTT CATCGATCGT ACAGATTTAC CAGCACACTA | 500 |
| 45 | GATAGAAATC GAGACCGACG CGGGCAGATC AATAGATTT TCTAGACGTT | 550 |
| | TTATTGGATC GTGAGATGAT TGATTGGGGT GGCGTGTGCA TACGATAGCG | 600 |
| 50 | GTGCACCGCC GATGTATCGG GGCATGTGCA CGTGGTTGGG TCTCAGCAGA | 650 |
| | CATATCACTA GACTGGTATC GTAATTACT AGTACTACTG GAAAGAGGAC | 700 |
| | TAAAAAGGCT AGGCCAAGTG CACGCATGTT GGGAACGTTG TTAAATTGAT | 750 |
| 55 | GAGTTTGTCC TTTGCTTGGG CTGGTATTAT TACCAAAAAA TGGTGTAGT | 800 |

| | | |
|----|--------------------------------------------------------|------|
| | CCCTGTACTT ATTAATGGGA AAATCTTAAC ATGACACTGG GGTTTATGAG | 850 |
| | TCTCCAATTG TATATTCTCA GCACTCAACT GATTTACTG ATACTGTAGT | 900 |
| 5 | GGAAATGACA CGTGAGCACC CCCCTCAAG GAATGCAATG CTTCTTCTG | 950 |
| | TTTTATATTA CAGGAACTAG AAGGAGCTTC CACCTTGAG TACAGAAGTA | 1000 |
| 10 | CTCCCCTCGT TCCAAAATAG ATGACTCAAC TTTGTACTAA TTTTGTACTA | 1050 |
| | TAGTTAGTAC AAAGTTGAGT CATCTATTT AGAACGGAGG GAGTAGTATC | 1100 |
| | GAAATTGAAG ACCCTGTAT TACTGTCTTG TTTTCAATG AAAATGGGAG | 1150 |
| 15 | GCCCATGCAG TAAGTCACAT GGGCACCTGG GAGGCTGGGA TCATGTGTGC | 1200 |
| | TTTGCAGAGT ACTAGACCCA GCTCACCCCTC TGTTAGATTA CTTGTTGGC | 1250 |
| 20 | TGCTACTTTG TGTTGCTGT GCAGTATATC AGACATCCTG AATTGGCAT | 1300 |
| | CTAGCTGAGA ACAGAATGCA GGTTGCACCA TTCTTATTAT TGCTAAACTG | 1350 |
| | TTGTCACGCA ATTATAAAAG AATGTGATCT TCTGAGTATT AATTAATCAT | 1400 |
| 25 | GTTCTGCTAA TATCTGTCCT CGCTCTGGTG TTGACAAATA TACCATATGA | 1450 |
| | ATATTTCNA TTTGCAACC AGGGATTGCT GAGGATTCCA TCGACAGCAT | 1500 |
| 30 | AATCGTGGCT GCAAGTGAGC AGGATTCTGA GATCATGGAT GCGAATGAGC | 1550 |
| | AACCTCAAGC TAAAGTTACA CGTAGCATCG TGTTGTGAC TGGTGAAGCT | 1600 |
| | GCTCCTTATG CAAAGTCAGG GGGGCTGGGA GATGTTGTG GTTCGTTACC | 1650 |
| 35 | AATTGCTCTT GCTGCTCGTG GTCACCGTGT GATGGTTGTA ATGCCAAGAT | 1700 |
| | ACTTGAATGG GTCCTCTGAT AAAAACTATG CAAAGGCATT ATACACTGCG | 1750 |
| 40 | AAGCACATTA AGATTCCATG CTTGGGGGA TCACATGAAG TGACCTTTT | 1800 |
| | TCATGAGTAT AGAGACAACG TCGATTGGGT GGGTACACAA TCACCTCTT | 1850 |
| | ATTCTCTGTT GAATTGTAGC AACTGTTAT CCTTGTAC ACTTCTTITA | 1900 |
| 45 | GCCCTGCAA GACATATGTG ATTTCCATAC TTTTGTGTTA TTTCCCTGT | 1950 |
| | ACTCTTGCTC ATGAAGGTCA AAATATCATA TATCCATGGA AGTCATGCAT | 2000 |
| 50 | GTGCCTAGTA TTTTGGTGT CGGTGCCTT AACCTTCAGG GATTAATACG | 2050 |
| | TGGAATTGTA TAACTAAAGT TTATTTATT GAAAAAAATT GTAGGTTGG | 2100 |
| | TGAGCCCACA GCCACGCAGT GGCACCACTG CTTGCACATG ATTTGCATT | 2150 |
| 55 | TCTGTTGCA CCGAGCACTT CATGTGAATA AGGTGTAAAA TCATAAAGTA | 2200 |

- 89 -

| | | |
|----|---------------------------------------------------------|------|
| | CCAATTTAT TCTGCCAATT GCAC TTAAGA GTATATA CAT TTATCTTGGC | 2250 |
| | CTCAATCATG GGAGTACTGT GCATTCAGTG CACCATCATT GTTCTAAGGA | 2300 |
| 5 | GAAAATGTGG GTGCAAGGAA GACACTTTG TCCCTTAATA AAAGGCAGGC | 2350 |
| | ACTCTGTTGT CATATAGATA GAAAGCAACA AACTTATTTC AAAGAGCTAA | 2400 |
| 10 | CAATGGCAA AGAACCAAAA AAAGCATGCT AAGGCCGTGA CACCAAAAGG | 2450 |
| | TGAGGGGGGC CTTGTGACTG ACAGCACCCC AACTATTGC CATTGTTTA | 2500 |
| | CTAAATGAAG ATCATTAG AAGCTCTCAG GAACTTCGAA AACAGTGGCT | 2550 |
| 15 | TTCCGTCCAC AGATCGTCTG TTAATATTGT TGTCAGTGA TACTTTTTT | 2600 |
| | GCTCCTTACA AGAGTGCCTA TGTTGACATA TACATTGTTA AGTTGTTCAT | 2650 |
| 20 | AAGTTTACTT CTTATTCTAA ACAGCAAGTG CCTAATGCTT GCATTTATT | 2700 |
| | TGGCTATTITA TTTTATTCT CATTCAATC AACACTTTG TTCAGGTGTT | 2750 |
| | TGTCGATCAT CCGTCATATC ATAGACCAGG AAGTTTATAT GGAGATAATT | 2800 |
| 25 | TTGGTGCTTT TGGTGATAAT CAGGTACACT AACTATACT AAGCTCCTAG | 2850 |
| | TTGACTAAGT CGTAAGTTGT ACCTCCTCGC TGACCGGCTG CTCTATGTCG | 2900 |
| | TGCAGTTCAAG ATACACACTC CTTGCTATG CTGCATGCGA GGCCCCACTA | 2950 |
| 30 | ATCCTTGAAT TGGGAGGATA TATTATGGA CAGAATTGCA TGTTGTTGT | 3000 |
| | GAACGATTGG CATGCCAGCC TTGTGCCAGT GTACGTTGTT TGTGGATCTG | 3050 |
| 35 | AAAGTCCAAT CTTTATTCA TTCTCTGCTT TGCA GTGTGC CCATGTCTAC | 3100 |
| | ATTTCTTTA TGCTTTTTC ATGTCTGTT TCATATTGCA TATATGCTTA | 3150 |
| | TGGAGTCTAA AAGTTACCGG AGGGAAATAAC TCTTAAGGAT TTCTCAATC | 3200 |
| 40 | AATTATCTTT AGCTTTAGTT AACATTACT GTGGCAAACA TAATGTGTTT | 3250 |
| | TGAGATTAC AAGTTCAGAG ATTGCACTTC ACTAGTTCGT AGCTAATCTG | 3300 |
| 45 | ATGTTTCCC CGAGAAAATG CCTAAAGCTT TGTGTCTTGA TGCATTGATA | 3350 |
| | GAAAAAGAGT TTATGTACAC TCCCAAAGAG GGGACCCAAA ATTACAACAC | 3400 |
| | CACACCCCTG AGAACTAGGC GCTGCCGGAA GAAGCGATGC AAGCCCCACT | 3450 |
| 50 | GCCCCCTGCCT TAGCTCAAAG CCGGGCGTCA GCTTGATTGT GTCAAGTAAG | 3500 |
| | CTAGCAGTGC TAGATTGCGC AAGGTCGATT CGTCGAAGAT GACAGTGTG | 3550 |
| 55 | CGCTGCTTCC AAATCCACCA AACTATGAGC ATGATCACTG GAGAAGTACC | 3600 |

- 90 -

| | | |
|----|----------------------------------------------------------|------|
| | TTTCTCGCG GCTGAGGGGG TGGACTGGTG GTCTGCTGCT GCCAGTTTC | 3650 |
| | AGATAATCTG AAAAATGCAT GTTTGATGA TTTTAGTATC TTGCGGACCC | 3700 |
| 5 | TGGGTACCAC CTAAGCTTTC ACACAGTAAT TTGCAGTTAC ACCTATAAAA | 3750 |
| | GTAACGGTCA TGATATGCAT GTGTTTGGG TAGATCATGG TGCATGCATT | 3800 |
| 10 | TTAGGAATTAA GGACATGCCA GAACCACGTG AGGCTTATGG GGCAATTCAT | 3850 |
| | TTGTTCCATT ATACGAGTCA TGAATATGGT TCAGCATGTT TGGACGCTAC | 3900 |
| | TTGTTGGGG CAATTCAGA TGGTGAATTG TAGCTGCTTG ATGTTGGCTA | 3950 |
| 15 | GCTGGCTTAT TTTGTACAAG TATCGATGTT AGATGCATAT TTCCTTTGT | 4000 |
| | TCTTGTGCTG TTTGCCATGT TGTATTCCCC TTTCTGCTG CCAGTGTGCG | 4050 |
| 20 | ATGTTAAATT GGTTTCATT ACATAATCAA CTTTGTGCT GACATCAGTC | 4100 |
| | ATTTTTATTTC AGCCTTCTTG CTGCAAAATA TAGACCATAC GGTGTTACA | 4150 |
| | GAGATTCCCG CAGCACCCCTT GTTATACATA ATTTAGCACA TCAGGTTGG | 4200 |
| 25 | GTCTATCACC TTTCATTATC CGTACATGGC TTTGTAAGTC GGTCACACG | 4250 |
| | TATCGTCATA CTGTATGTTA TTTCAATGTC ATTAGGGTGT GGAGCCTGCA | 4300 |
| 30 | AGTACATATC CTGATCTGGG ATTGCCACCT GAATGGTATG GAGCTTACA | 4350 |
| | ATGGGTATTT CCAGAATGGG CAAGGAGGCA TGCCCTTGAC AAGGGTGAGG | 4400 |
| | CAGTTAACCTT TTTGAAAGGA GCAGTTGTGA CACCGAGATCG AATTGTGACC | 4450 |
| 35 | GTCAGTCAGG TGAAATACTC AATACTTCTC TTTTTCTTT GCGGGATGTT | 4500 |
| | CTTCAGTTCA ATTGCCCTGT CTTTCACCCA ATTAAGAAAT GATTTAATCT | 4550 |
| 40 | TTTGTTCATA GGGTTATTCA TGGGAGGTCA CAACTGCTGA AGGTGGACAG | 4600 |
| | GGCCTCAATG AGCTCTTAAG CTCCCGAAAA AGTGTATTGA ATGGTAACCA | 4650 |
| | TATTTGAATC CACTTATCTT CTCTGAAAC ATATTTACAG AAATAGATGG | 4700 |
| 45 | ATGGGTTGCA AGAATAAATT CAGTTGCTC TTTCGGTATG AAGGAATTGT | 4750 |
| | AAATGGAATT GACATTAATG ATTGGAACCC CACCACAGAC AAGTGTCTCC | 4800 |
| 50 | CTCATCATTAA TTCTGTCGAT GACCTCTCTG GAAAGGTGTG TGGATAGTAC | 4850 |
| | CCTATATAAT AACATGTATA TCTGATCTAG TACTTTCTTT TTCTTTGCTA | 4900 |
| | GTTCGCTTCC CATGATGTTC TCACTAACTA ATCCTATGTG GTTGGCATA | 4950 |
| 55 | CTTGTCAAGGC CAAATGTAAA GCTGAATTGC AGAAGGAGCT GGGTTACCT | 5000 |

00000000000000000000000000000000

- 91 -

| | | |
|----|---------------------------------------------------------|------|
| | GTAAGGGAGG ATGTCCTCT GGTTAGATAC AAACCCCTAA GATATATATT | 5050 |
| | TTTTAAATCC CTAAAAAAA CTTGCCGATC ATCTCATTAG CTTGATTAC | 5100 |
| 5 | AGATTGGCTT TATTGGAAGA CTGGATTACC AGAAAGGCAT TGATCTCATT | 5150 |
| | AAAATGGCCA TTCCAGAGCT CATGAGGGAG GACGTGCAGT TTGTAAGTTC | 5200 |
| | ATATTCTTT TCTTGAGACT AGAGTATAAA TCAAACATGT AGGTGTGGGG | 5250 |
| 10 | TGGTATAATA CAGACATAAG TTCCAGCTAT TGCTTCCATG AGAATTAA | 5300 |
| | TGCTATTCAAG TAATATGCTA CTGCAAGTT TGAAACAAAG TTGGAAGCAA | 5350 |
| 15 | TAAATATATG TGTAGCACTG ACCATGCAGT GCCACTATAG CTGGAATGTC | 5400 |
| | CTGTAGTCTA TGTGATCTAA CACACTAAC AACATGTTT CGCATAACAAA | 5450 |
| 20 | CACATGCGTG CGCGCAACAA ACATACTCTA CAATAAAATT GGCTTGGTGA | 5500 |
| | ACTGCAGACA TGCTCTTATC TCCATTCAA CATTCTTGT TTCAACATTG | 5550 |
| | GCTGAAGACT AAGAGAAGGG GGACCCAGGG TGATGTAGCC AACTAGATCC | 5600 |
| 25 | AGTAAGGAAG CTAGCCGAGC CTAGGAGGAT TCGCTTAGGT AGCTGGAACG | 5650 |
| | TAGGGTCTCT GACAGGGAAG CTTCGGGAGC TAGTCGATGC AGTGGTGAGG | 5700 |
| | AGAGGTGTTG ATATCCTTG CGTCCAAGAA ACCAAATGTA GGGGACAGAA | 5750 |
| 30 | GGCGAAGGAG GTGGAGGATA CCGGCTTCAA GCTGTGGTAC ATGGGACGGC | 5800 |
| | TGCAAACAGA AATGGCGTAG GCATCTTGAT CAACAAGAGC CTTAAGTATG | 5850 |
| 35 | GAGTGGTAGA CGTCAAGAGA CGTGGGGACC GGATTATCCT CGTCAAGCTG | 5900 |
| | GTAGTTGGGG ACTTAGTTCT CAATGTTATC AGCGTGTATG CCCCCGCAAGT | 5950 |
| 40 | AGGCCACAAT GAGAACGCCA AGAGGGAGTT CTGGGAAGGC CTGGAAGACA | 6000 |
| | TGGTTAGGAG TGTACCGATT GGCGAGAACG TCTTCATAGG AGGAGACCTC | 6050 |
| | AATGGCCACG TGGGTACATC TAACATAGGT TTTGAAGGGG CACATGGGG | 6100 |
| 45 | CTTTGGCTAT GGCATCAAGA ATCAAGAAGA AGATGTCTTA CGCTTGCTC | 6150 |
| | TAGCCTACGA CATGATTGTA GCTAACACCC TCTTAGAAA GAGAGAATCA | 6200 |
| | CATCTGGTGA CTTTAGTAG TGGCCAACAC TAGCCAGATC GATTTCATCC | 6250 |
| 50 | TCTCGAGAAG AGAAGATAGG TGTGCGCGCC TAGACTGCAA GGTGATACCT | 6300 |
| | TCGGATTCTGT GTCCAGCGGG ATAAGCGTGC CAAAGTCGCT AGAATGAAGT | 6350 |
| 55 | GGTGGAAAGCT CAAGGGGGAG GTAGCTCAGG CGTTCAAGGA GAGGGTCATT | 6400 |

50000 4250000000

- 92 -

| | | |
|----|---------------------------------------------------------|------|
| | ACGGAGGGCC CTTGGGAGGA AGGAGGGAT GCGGACAATG TGTGGATGAA | 6450 |
| | GATGGCGACT TGCATTGTA AGGTGGCCTC GGAGGAGTGT GGAGTGTCCA | 6500 |
| 5 | GGGGATGGAG AAGCGAAGAT AAGGATAACCT GGTGGTGGAA TGATGATGTC | 7000 |
| | CAGAAGGCAA TTAAAGAGAA GAAAGATTGC TTTAGACGCC TATACTTGGA | 7050 |
| 10 | TAGGAGTGCA GTCAACATAG AAAAGTACAA GATGGCGAAG AAGGCCGCAA | 7100 |
| | AGCGAGCTGT CAGTGAAGCA AGGGGTCGGG CATATGAGGA TCTCTACCAA | 7150 |
| | CGGTTAGGCA CGAAGGAAGG CGAAAGGGAC ATCTATAAGA TGGCCAAGAT | 7200 |
| | CCGAGAGAGA GGAAGACGAG GGATATTGGC CAAGTCAAAT GCATCAAGGA | 7250 |
| 15 | TGGAGCAGAC CAACTCTTGG TGAAGGACGA GGAGATTAAG CATAGATGGC | 7300 |
| | GGGAGTACTT CGACAAGCTG TTCAATGGGG AGGATGAGAG TCCTACCATT | 7350 |
| | GAACTTGACG ACTCCTTGA TGAGACCATC ATGCGTTTA TGCAGCGAAT | 7400 |
| | CCAGGAGTCC GAGGTCAAGG AGGCTTAAA AAGGAGGCAA GGCGATGGC | 7450 |
| | CCTGATTGTA TCCCCATTGA GGTGTGGAAA GGCCTCGGGG ACATAGCGAT | 7500 |
| 20 | AGTATGGCTA ACCAACGCTAT TCAACCTCAT TTTTCGGCA AACAAAGATGC | 7550 |
| | CAGAAGAATG GAGACGAAGT ATATTAGTAC CAATCATCAA ACAGGGGGGA | 7600 |
| | TGTCAGAGT TGTACTAATT ACCATGGAATTAAAGCTGATG AGCCATACAA | 7650 |
| | TGAAGCTATG GGAGAGAATC ATTGAGCACC GCTTAAGAAG AATGACAAGC | 7700 |
| | GTGACCAAAA ATCAGTTGG TTTCATGCCT GGGAGGTCGA CCATGGAAAC | 7750 |
| 25 | CATTTCTTG GTACGACAAC TTATGGAGAG ATACAGGGAG CAAAAGAAGG | 7800 |
| | ACTTGCATAT GGTGTCATT GACTTGAAGA AGGCCTATAA TAAGATACCG | 7850 |
| | CGGAATGTCA TGTGGTGGGC CTTGGAGAAA CACAAAGTCC CAGCAAAGTA | 7900 |
| | CATTACCTC ATCAAGGACA TGTACGATAA TGTTGTGACA AGTGTTCGAA | 7950 |
| | CAAGTGATGT CGACACTAAT GACTTCCGA TTAAGATAGG ACTGCATCAG | 8000 |
| 30 | GGGTCAGCTT TGAGCCCTTA TCTTTTGCC TTGGTGATGG ATGAGGTCAC | 8050 |
| | AAGGGATATA CAAGGAGATA TCCCAGGGTG TATGCTCTT GTGGATGATT | 8100 |
| | TGGTGCTAGT TGACGATAGT CGGGCGGGGG TAAATAACAA GTTAGAGTTA | 8150 |
| | TGGAGACAAA CCTTGGAAATC GAAAGGGTTT AGGCTTAGTA GAACTAAAAC | 8200 |
| | CGAGTACATG ATGTGCGGTT TCAGTACTAC TAGGTGTGAG GAGGAGGAGG | 8250 |

00000000000000000000000000000000

- 93 -

| | | |
|----|----------------------------------------------------------|------|
| | TTAGCCTTGA TGGGCAGGTG GTACCCCAGA AGGACACCTT TCGATATTG | 8300 |
| | GGGTCAATGC TGCAGGAGGA TGGGGTATT GATGAAGATG TGAACCATCG | 8350 |
| | AATCAAAGCT GGATGGATGA AGTGGGCCA AGCTTCTGGC ATTCTTGTG | 8400 |
| | ACAAGAGAGT GCCACAAAAG CTAAGGCAAG TTCTACAGGA CGGCGGTTCG | 8450 |
| 5 | ACCCGCAATG TTGTATGGCG CTGAGTGTG GCCGACTAAA AGGCGACATG | 8500 |
| | TTCAACAGTT AGGTGTGGCG GAGATGCGTA TGTTGAGATG GATGTGTGGC | 8550 |
| | CACACGAGGA AGRATCGAGT CCGGAATGAT GATATACGAG ATAGAGTTGG | 8600 |
| | GGTAGCACCA ATTGAAGAGA AGCTTGTCCA ACATCGTCTG AGATGGTTG | 8650 |
| | GGCATATTCA GCGCACGCCCT CCGAAAACCTC CAGTGCATAA CGGACGGCTA | 8700 |
| 10 | AAGCGTGCAGG AGAATGTCAA GAGAGGGCGG GGTAGACCGA ATTTGACATG | 8750 |
| | GGAGGAGTCC GTTAAGAGAG ACCTGAAGGT TTGGAGTATT ACGAAAGAAC | 8800 |
| | TAGCTATGGA CARGGGTGCG TGGAAGCTTG TTATCCATGT GCCAGAGCCA | 8850 |
| | TGAGTTGATC ACGAGATCTT ATGGGTTCA CCTCTAGCCT ACCCCAACTT | 8900 |
| | GTTTGGGACT AAAGGCTTTG TTGTTGTTGT TGTTGTTGTT GTTGTAGCCA | 8950 |
| 15 | ACTAAATCCA GTTGATCAGT GGTTTTACT CTTATTTTA CAGGTCAATGC | 9000 |
| | TTGGATCTGG GGATCCAATT TTTGAAGGCT GGATGAGATC TACCGAGTCG | 9050 |
| | AGTTACAAGG ATAAATTCCG TGGATGGGTT GGATTTAGTG TTCCAGTTTC | 9100 |
| | CCACAGAATA ACTGCAGGGT ATGCCGAGAA CTTCTTAACA AGACCTTCGT | 9150 |
| | TATCAGCTTG GATATATTAT AATGTTCAAA ACATTTATGT CTCTTTTT | 9200 |
| 20 | GTGCAGTTGC GATATATTGT TAATGCCATC CAGGTTGAA CCTTGTGGTC | 9250 |
| | TTAACATCAGCT ATATGCTATG CAATATGGTA CAGTCCTGT AGTCATGGA | 9300 |
| | ACTGGGGGCC TCCGAGTAAG ACAACTGCCT TGAAAATTAT CGTTATTTG | 9350 |
| | GCTCCAACGC AAATGTTCTA ATTGGCTCGT GTATTCAACA GGACACAGTC | 9400 |
| | GAGACCTTCA ACCCTTTGG TGCAAAAGGA GAGGAGGGTA CAGGGTACGC | 9450 |
| 25 | ACTGCTCAAT TTTCAGCTAAC TTTCAGTTA TCTTTTGCA ATGTCTTGGG | 9500 |
| | GGTTCATTGC GCCATAAAC AACTTGTGAT AATTAACTGT TACTGTTCTG | 9550 |
| | TACTTGCAGG TGGCGTTCT CACCGCTAAC CGTGGACAAG ATGTTGTGGG | 9600 |
| | TAAGTTTTG CTGAGCTCTT GTCCGGTTAT AGGATCGACC TTGGCTGTAG | 9650 |

DOCUMENT RELEASED BY GOVERNMENT OF CANADA

- 94 -

| | | |
|----|---------------------------------------------------------|-------|
| | CATGGTACCT TAGTCCCCCT TGTATATAGA CCTAACCTGA TGGACTCACT | 9700 |
| | TTGTCTACAC TAATCATAGT AGTCGATTGC CCGGAGGCCT TTTGCTTGAA | 9750 |
| | TTCTGCTAAT TTAATTTCA TGACCGATAAC TCATACCATG GTTTGGTTCT | 9800 |
| | CCGATGGGG CCAGAATGGC GTCTAGTGTG TGCGATCTGT GTAACTAGCC | 9850 |
| 5 | AATGCCGGGT TGTTCCAAGT GAAAATTAC CTTTGACCA TTGTGCAGGC | 9900 |
| | ATTGCGAACCGCGATGTCGA CATTCAAGGGA GCACAAGCCG TCCTGGGAGG | 9950 |
| | GGCTCATGAA GCGAGGCATG ACGAAAGACC ATACGTGGGA CCATGCCGCC | 10000 |
| | GAGCAGTACG ACCAGATCTT CGAATGGGCC TTCGTGGACC AACCTACGT | 10050 |
| | CATGTAGACG GGGACTGGGG AGGTGAAAGC GCGGGTCTCC TTGAGCTCTG | 10100 |
| 10 | AAGACATGTT CCTCATCCTT CCGCGGCCCCG GAAGGATACC CCTGTACATT | 10150 |
| | GCGTTGTCCT GCTACAGTAG AGTCGCAATG CGCCTGCTTG CTTGGTCCGC | 10200 |
| | CGGTTCGAGA GTAGATGACG GCTGTGCTGC TGCGGGCGTG ACAGCTTCGG | 10250 |
| | GTGGATGACA GTTACAGTTT TGGGAAATAA GGAAGGGATG TGCTGCAGGA | 10300 |
| | TGGTTAACAG CAAAGCACCA CTCAGATGGC AGCCTCTCTG TCCGTGTTAC | 10350 |
| 15 | AGCTGAAATC AGAAACCAAC TGGTGAATCT TTAGCCTTAG CGATTGTGAA | 10400 |
| | GTTTGTGCA TTCTGTGTAT GTTGTCTTGT CCTTAGCTGA CAAATATTTG | 10450 |
| | ACCTGTTGGA TAATTCTATC TTTGCTGCTG TTTTCTTT GGTCAAAAGA | 10500 |
| | GGGGTTCCCT CCGATTTCAT TAACGAAACC ACCAAAATAA CAGCACCCAG | 10550 |
| | TGCAGGTCTC AGGTTCAAGAT ATACTTAAGA CTACTAAATC TAACAGCAGC | 10600 |
| 20 | TAAAAAGCTT AAAGATTCAG GCGACATAAC CGAACAAAAT CCACAACCGA | 10650 |
| | AGGGACCAAA GCAGGACAAG TAAAAAGGCA GNCGACACAA AGCGCAGGTC | 10700 |
| | GCTGAAAAGG CAAGCAGACA GAGGTCTGCA TTCTGTCAAC ACCACTTGTG | 10750 |
| | AAAAATGAAG AGAAGATCGA GAATTCCCGG GAATCCG | 10787 |
| 25 | (2) INFORMATION FOR SEQ ID NO: 14: | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 647 amino acids | |
| | (B) TYPE: amino acid | |
| | (C) STRANDEDNESS: single | |
| | (D) TOPOLOGY: linear | |
| 30 | (ii) MOLECULE TYPE: protein | |
| | (iii) HYPOTHETICAL: NO | |

- 95 -

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *triticum tauschii*
(F) TISSUE TYPE: Endosperm

5 (ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..647
(D) OTHER INFORMATION:/product= "deduced amino acid sequence for SSS I"

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

| | |
|----|------------------------------------------------------------------------------------------------------------------------------------------------|
| 15 | Met Ala Ala Thr Gly Val Gly Ala Gly Cys Leu Ala Pro Ser Val Arg 1 5 10 15 |
| 20 | Leu Arg Ala Asp Pro Ala Thr Ala Ala Arg Ala Ser Ala Cys Val Val 20 25 30 |
| 25 | Arg Ala Arg Leu Arg Arg Leu Ala Arg Gly Arg Tyr Val Ala Glu Leu 35 40 45 |
| 30 | Ser Arg Glu Gly Pro Ala Ala Arg Pro Ala Gln Gln Gln Gln Leu Ala 50 55 60 |
| 35 | Pro Pro Leu Val Pro Gly Phe Leu Ala Pro Pro Pro Pro Ala Pro Ala 65 70 75 80 |
| 40 | Gln Ser Pro Ala Pro Thr Gln Pro Pro Leu Pro Asp Ala Gly Val Gly 85 90 95 |
| 45 | Glu Leu Ala Pro Asp Leu Leu Leu Glu Gly Ile Ala Glu Asp Ser Ile 100 105 110 |
| 50 | Asp Ser Ile Ile Val Ala Ala Ser Glu Gln Asp Ser Glu Ile Met Asp 115 120 125 |
| 55 | Ala Asn Glu Gln Pro Gln Ala Lys Val Thr Arg Ser Ile Val Phe Val 130 135 140 |
| 60 | Thr Gly Glu Ala Ala Pro Tyr Ala Lys Ser Gly Gly Leu Gly Asp Val 145 150 155 160 |
| 65 | Cys Gly Ser Leu Pro Ile Ala Leu Ala Ala Arg Gly His Arg Val Met 165 170 175 |
| 70 | Val Val Met Pro Arg Tyr Leu Asn Gly Ser Ser Asp Lys Asn Tyr Ala 180 185 190 |
| 75 | Lys Ala Leu Tyr Thr Gly Lys His Ile Lys Ile Pro Cys Phe Gly Gly 195 200 205 |
| 80 | Ser His Glu Val Thr Phe Phe His Glu Tyr Arg Asp Asn Val Asp Trp 210 215 220 |
| 85 | Val Phe Val Asp His Pro Ser Tyr His Arg Pro Gly Ser Leu Tyr Gly 225 230 235 240 |
| 90 | Asp Asn Phe Gly Ala Phe Gly Asp Asn Gln Phe Arg Tyr Thr Leu Leu 245 250 255 |
| 95 | Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile Leu Glu Leu Gly Gly Tyr 260 265 270 |

Ile Tyr Gly Gln Asn Cys Met Phe Val Val Asn Asp Trp His Ala Ser
 275 280 285

5 Leu Val Pro Val Leu Leu Ala Ala Lys Tyr Arg Pro Tyr Gly Val Tyr
 290 295 300

Arg Asp Ser Arg Ser Thr Leu Val Ile His Asn Leu Ala His Gln Gly
 305 310 315 320

10 Leu Glu Pro Ala Ser Thr Tyr Pro Asp Leu Gly Leu Pro Pro Glu Trp
 325 330 335

Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu Trp Ala Arg Arg His Ala
 340 345 350

Leu Asp Lys Gly Glu Ala Val Asn Phe Leu Lys Gly Ala Val Val Thr
 355 360 365

20 Ala Asp Arg Ile Val Thr Val Ser Gln Gly Tyr Ser Trp Glu Val Thr
 370 375 380

Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu Leu Leu Ser Ser Arg Lys
 385 390 395 400

25 Ser Val Leu Asn Gly Ile Val Asn Gly Ile Asp Ile Asn Asp Trp Asn
 405 410 415

30 Pro Thr Thr Asp Lys Cys Leu Pro His His Tyr Ser Val Asp Asp Leu
 420 425 430

Ser Gly Lys Ala Lys Cys Lys Ala Glu Leu Gln Lys Glu Leu Gly Leu
 435 440 445

35 Pro Val Arg Glu Asp Val Pro Leu Ile Gly Phe Ile Gly Arg Leu Asp
 450 455 460

Tyr Gln Lys Gly Ile Asp Leu Ile Lys Met Ala Ile Pro Glu Leu Met
 465 470 475 480

40 Arg Glu Asp Val Gln Phe Val Met Leu Gly Ser Gly Asp Pro Ile Phe
 485 490 495

45 Glu Gly Trp Met Arg Ser Thr Glu Ser Ser Tyr Lys Asp Lys Phe Arg
 500 505 510

Gly Trp Val Gly Phe Ser Val Pro Val Ser His Arg Ile Thr Ala Gly
 515 520 525

50 Cys Asp Ile Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly Leu Asn
 530 535 540

Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val Pro Val Val His Gly Thr
 545 550 555 560

55 Gly Gly Leu Arg Asp Thr Val Glu Thr Phe Asn Pro Phe Gly Ala Lys
 565 570 575

60 Gly Glu Glu Gly Thr Gly Trp Ala Phe Ser Pro Leu Thr Val Asp Lys
 580 585 590

Met Leu Trp Ala Leu Arg Thr Ala Met Ser Thr Phe Arg Glu His Lys
 595 600 605

- 97 -

| | | | | | | | | | | | | | | | | |
|--|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | Pro | Ser | Trp | Glu | Gly | Leu | Met | Lys | Arg | Gly | Met | Thr | Lys | Asp | His | Thr |
| | 610 | | | | | | 615 | | | | | | | | | 620 |

| | | | | | | | | | | | | | | | | |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 5 | Trp | Asp | His | Ala | Ala | Glu | Gln | Tyr | Glu | Gln | Ile | Phe | Glu | Trp | Ala | Phe |
| | 625 | | | | | 630 | | | | | 635 | | | | | 640 |

| | | | | | | | | | | | | | | | | |
|--|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|--|--|--|--|--|
| | Val | Asp | Gln | Pro | Tyr | Val | Met | | | | | | | | | |
| | | | | | | 645 | | | | | | | | | | |

10

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5072 base pairs

(B) TYPE: nucleic acid

15

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: triticum tauschii

(F) TISSUE TYPE: Endosperm

25

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 1..4993

(D) OTHER INFORMATION:/function= "region containing
promoter of SSS I"

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|----|
| TCTAGATGCA | TGCTGGATAG | CGGTCGATGT | GTGGAGTAAT | AGTAGTAGAT | GCAGAACCGT | 60 |
|------------|------------|------------|------------|------------|------------|----|

35

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| TTCGGTCTAC | TTGTCGCGGA | CGTGATGCCT | ATATACTGAA | TCATACCTAG | ATATTCTCAT | 120 |
|------------|------------|------------|------------|------------|------------|-----|

| | | | | | | |
|------------|------------|------------|-------------|------------|------------|-----|
| AACTATGCTC | AATTCTATCA | ATTGCTCGAC | AGTAATTGCGT | TTACCCACCG | TAATACTTAT | 180 |
|------------|------------|------------|-------------|------------|------------|-----|

40

| | | | | | | |
|------------|------------|------------|------------|-----------|------------|-----|
| GATCTTGAGA | GAAGTCACTA | GTGAAACCTA | TGCCCCCGAG | GTCTATTTG | CATCATATTA | 240 |
|------------|------------|------------|------------|-----------|------------|-----|

| | | | | | | |
|------------|------------|------------|-----------|------------|------------|-----|
| ATCTTCCAAT | ACTTAGTTAT | TTCCATTGCC | GTTTATTTA | CTTTGTATCT | TTATTTCTTT | 300 |
|------------|------------|------------|-----------|------------|------------|-----|

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| TTATTATAAA | AAATACCAAA | AATATTATCT | TATCATATCT | ATCAGATCTC | ATTCTCGTAA | 360 |
|------------|------------|------------|------------|------------|------------|-----|

45

| | | | | | | |
|------------|------------|------------|------------|-----------|----------|-----|
| GTGACCGTGA | AGGGATTGAC | AACCCCTTTA | TCGTGTTGGT | TGCGAGGTT | TTGTTTGT | 420 |
|------------|------------|------------|------------|-----------|----------|-----|

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| GTGTAGGTGC | GTGTGACTCG | CACGTCTCCT | ACTGGATTGA | TACCTTGGGT | TTTCAAAAAC | 480 |
|------------|------------|------------|------------|------------|------------|-----|

50

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| TGAGAAAAAT | ACTTACGCTA | CTTTACTGCA | TAACCCTTTC | CTCTTTAAAA | AAAAAAACCA | 540 |
|------------|------------|------------|------------|------------|------------|-----|

| | | | | | | |
|------------|------------|------------|------------|------------|-----------|-----|
| ACGTAGTATT | CAAGAGGTAG | CACGCTACCA | TCCTCTCCAA | CAGGAGCGCG | GAGATCTTG | 600 |
|------------|------------|------------|------------|------------|-----------|-----|

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| TCCGGCAGGT | TGATGCGGGC | CGGGGAAGAA | CTCCAGCTGC | CTTGGCCAGC | TTGGTCGTGA | 660 |
|------------|------------|------------|------------|------------|------------|-----|

55

| | | | | | | |
|------------|------------|------------|------------|------------|-----------|-----|
| GCCGCCCCAG | CGGCGTCTTG | AACCTGTCCA | CGTAGCGCTC | CCTGACACGC | GGCGTGAAC | 720 |
|------------|------------|------------|------------|------------|-----------|-----|

| | | | | | | |
|------------|-------------|------------|------------|------------|------------|-----|
| GAGAAGGCTT | GTGCGATGAAC | TCCAGCTGTT | GTGCCAGCCT | AGCTTGCGCC | TTCTTCTGCT | 780 |
|------------|-------------|------------|------------|------------|------------|-----|

60

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| GGGTCATGCC | CTTCGAGAAA | CCCACCTTGG | CCACCCCTGT | GCTTGAGCGG | CGCGCCACCT | 840 |
|------------|------------|------------|------------|------------|------------|-----|

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| CAGCAGGCCG | CGGCGTGGGG | ATGAAGAGGG | TGTCTGCTTC | CGGAGCAGGC | GGGTCGGCGT | 900 |
|------------|------------|------------|------------|------------|------------|-----|

DOCUMENT FINGERPRINT

- 98 -

| | | | | | | | |
|----|------------------|-----------------------------|--------------|------------|---------------------|------------|------|
| | TGAACCTGAA | AGGC GG TG GC CCC AT GAT GG | ATGGGGGGAG | CATGCCAAAG | ACTTGTTGA | 960 | |
| | GGAAAAGTGGT | GTTGGCGTCC | ACCTCCAGTG | CCTGCAGTTT | GGAAGCCAGA | CGATTGGCGT | 1020 |
| 5 | CGATCTCTGG | CTCCGGCTGG | AAGGAGGCTC | GACGCTCCGG | TGTGCCAGAA | CGCAAAGGGA | 1080 |
| | GGAGCGGCAG | CTCTGGCTGA | GCAGACCCCCG | CGCCCATGTA | CTCTGCATTG | GGCCAAGGCT | 1140 |
| 10 | GCAGGGGCAA | GCCACC GG GA | TGGGGGCGCG | AGGTGGACTG | CGCACCGGAG | GAAGGCCAAG | 1200 |
| | CTCAACCTCG | GTGAGGTTCG | CCCCAGACCA | GGCGGGCAGG | CTCGGGTCCA | CAAAGGGCCA | 1260 |
| | AACCGCCTCG | TCCGCCCGA | AACTGTCCAG | GACAGACGGC | GGACGACGGA | AGGCCGTGTC | 1320 |
| 15 | GTCGAGCTCG | AGCAGCAGAG | GGTCCGTGCG | GGTGATGTCT | TGCCAAATGG | ACTCCACCTC | 1380 |
| | CAGCAGGAAG | GGGGACTGGT | CCATCGCCCC | TGGCCAAGCC | ACTGGTACGC | CAAAGATGGC | 1440 |
| 20 | ATCAGCAGCG | TTTGCACCAAG | GGGGAGCAGC | CACACCTTGG | AGGACAGGGA | GGGTGCGGAC | 1500 |
| | GTCGACGGCA | GCAAAACGTG | GCTGGAGCAA | GTTGCCGTG | CGTGCCGGCC | TCGGCGAGCG | 1560 |
| | CGAGCGGCTG | TAGGAGCGCT | CGGTGCCCTC | AGACTCGGAC | AGTGCGCCAG | TGGGAGAGCC | 1620 |
| 25 | ATGGCCACGC | CGGCCACCAC | TGGACGTGCC | ATGGCGCTGG | TCCTGACGGC | GCCTGGATGG | 1680 |
| | CCC GT C C T C G | CGGGCAGCTC | CACCTGAGCG | GCACCCGAGG | AGCACACCCC | GCCAAGCTGG | 1740 |
| 30 | GCCAGGGCGG | CTGCGGCAC | GGCGACGGCC | GCGGTCGCGG | TCTGCACCAT | CATCTTCATC | 1800 |
| | TTCGT CAT CG | TGGCGCCTCG | GACAAGGATG | CTCGCTGTCA | CCGACGCGAG | GGACGTGAGC | 1860 |
| | CGGCTCAGCC | CGCCCTTCCT | CGACGTGGCG | AGCCCTGCGG | ATATGCTCCT | CGAGCGGCCA | 1920 |
| 35 | TTGGGGGTG | TTGGCGCGCG | GCATCTCGGG | GTCGCGGTCA | GCTATCGGGG | TGTAGTCCTT | 1980 |
| | TGTGGTGTCC | AGGTGGATGA | GCAGAGAGAA | ATCCGGCCCC | TCTAGCCCC | CGTCCGGGG | 2040 |
| 40 | GCAGCCCTCC | GGCAGCGTCT | GGCGGCCCT | GGGGTCCAGG | GGTCGATCGA | TGATGGAGAA | 2100 |
| | CCCCCTTTG | GTGGGGATGT | CGTCCGGACT | CCATGCCAC | A C C C A G G C A A | AGAGGCAGGC | 2160 |
| | CGTGTGGAG | AGGGAGGTG | TCTGCCGCTC | CAACCAGTCG | ACGTGGCATG | TCTTCCCGAG | 2220 |
| 45 | CGCATCCTGC | CCCGCCTCCT | TGTTCCAGGA | CTGCACCGGC | ATGTTCTCGA | CGGCGATGCG | 2280 |
| | GCAGTAGTAC | CGCCAGACAC | GGCGGTGGCC | GTGTGCCGAT | GGTGACCAGG | CCGACAGGGA | 2340 |
| 50 | GAGCGCGACG | CCCCAGCAGG | AGACGACCCC | AGCGTCGAAA | GCGATGTCCC | GGTGCCTGAA | 2400 |
| | GTGGACGAGC | CCAGAGATGG | CCAGGCGCAT | TGACGCGGGG | AAGGGGAAGG | AGTTAGGATG | 2460 |
| | GGCGACCGGG | CCGGAGTGAA | CCGCGGCGTG | GTGGCCGACG | GGGCTGGAGA | GGCAGAGGCG | 2520 |
| 55 | GAGTCATCCG | AGAGAGGTGT | ATCAGTGGCT | CTGCACAATA | CCCAGTGTG | CCACATCATA | 2580 |
| | TCCTGCTGAA | TAACCACACA | TGTGTACTGT | CGTTAAATAA | ATCATTGGTC | ACGCGAACCC | 2640 |
| 60 | GGAAAAAGAC | GGCGAAAAAT | TCACGGACAC | ACGACTAGTA | GTACCCAATA | TACTCGGCAA | 2700 |
| | AAACACTGAC | ACGTCGTTTT | GC GTTGT CGG | CCGGTGTGT | CGAGTCATTG | TACTATGTT | 2760 |
| | TGTCGTTTCT | TTCTTTCTC | CAAATCGACA | AACC GTTGT | CTTGGTTAA | AAAACAGAAA | 2820 |
| 65 | CATACAAAAT | CAAATGAATG | CATTCAAGGG | CCGGTAATCC | AATTCTGAGC | CCAGGCTCAG | 2880 |
| | CTACACCCGC | CCTTACAAA | AAATCAAAT | AAATACTAGA | AAAATTCAA | AAATTCCAAT | 2940 |

- 99 -

TTGTTTGTGC GTGGTAGATA ATTTGATGCG TGAGGTACGC TTCAATTTC AAATTATTTG 3000
5 GACATCTGAG CAGCTCTCAG CAAAAAAGAC AAATTGGGG TCTGTAAAAA TGTTTACTGT 3060
TCATGCACTG TTCTGACCCG ATTTGTCTTT TTTGCTGAGA GCTTCTCAGA AGTCAAATG 3120
AGCTAAAATT TTGAGCGGAG CTTACGTGAT AAAATGTCTA TCATGCAAAA AAGGATTGGA 3180
10 ATTTTTGAA TTTTTTTTAT TTTTGAT TTGTTCCCTG GACGGGTGCA GATAAGCCTG 3240
GGCACCGAAA CGCCGCACTC AGGCTCATCC TTTTCTATAA AAGAAAAGAA ATACATACAA 3300
15 TTTCCCTCTG TTTTTGAGC AAGGGGCACC ACCCACAAA GAGTTTCAA CTCACATGGT 3360
ATTAGAGCAT CTACAGCCGG GCGTCTCAA CCAGCCTCAT ACGCTTGAGC GGGTCGCCCT 3420
GGTCACGATT TTTTGACCCA GACGGGCCCC TCAAACGGTC CTTAACGCC CAGGCTGACC 3480
20 GACAACCCAC ATATCCAGCC CAAATATGGG GTGGATATGG GGGCGCCCGG GCACGCCAGC 3540
CCGCGGACAC CACACATCTT CAGTTCTAA TTTGAGATAT CCGGATGTGG AATGCGTTT 3600
25 TGAGGGGTGA CCGGTCCCTG TCCGTGGATG CGCCCGGACG TTTGAGGGGT TGGATTGCC 3660
AAGTCTGATT AGAGATGCTC TTAGGTGTT CACCCCCATC CCTTGATGGC TAGGGCAAAC 3720
TCTCCCTCC AAACCTTGTC GGCGAGCCTG TGGATTCTTC TCTCCTCTGC CCGCTGCTCC 3780
30 GGC GGCTGAT GGC GGGGAGG AGAATCCCAG TGTCTCGCT TGGTTAGTTG TTTAAGTTAC 3840
GTACTTTTT AGTCCTCGCA GGTGCGGCGT TCGGACGTAT GGTCGTGCTT CTTTTTGAG 3900
35 TTTGCTTCC GGGCTCTGAT CCTCCTCGAG TTCGTCCATC TGGACGTACT CGACGGAGCT 3960
CCGGCATAGA TTCCATATCAT CGTCTTGGTG AGGTGAGGTT ATGGTTTCTT GTCATGTGGG 4020
CAGATTGGT GCCAGATGCT TCATATCTAT TCAAGGGTTC AGCGGCAACA ACTGCGGCTC 4080
40 CAGAGCGATG GTCCCTTAAGG GCACGTGCAC GAAGACTTCA CGGCTGTTAT CGACAAGGTC 4140
AAGCCGGCTC CGATAGGGGA GCAGCGACAG CGGCGCGTCA ACCGCTCGTT CTGGCGGCAG 4200
TAGTGGTCGT TCGGTGCTCT CGGAACCTCG ATGTAATTTC TATGATTTA GAGATGCTTT 4260
45 GTACTTCCGA TCGATGAACT CTGATAATAG ATATCTCTTC TCTCGAAAAA AAAGAGAGTT 4320
TTCAACTGAA AACAAAAGAG TTTCACTAGT TCTTCTTTA GAAACAGAGT TTCACTAGCA 4380
50 CTTTTTTTG CGAGAAGTCG AGTTCACTA AGTACTAAC CCACGCAATT ATTCTCAAAA 4440
AAAAAAACCA CGCAACTGTC TGGATCCATC TTCGTTTTT CCCCAGAAT CGTCTGGATC 4500
55 CATTTCGTG TCGGAGGCAT CCTCTCATTG TGACGGCCC AGCTCTCTTC TCGCCGGCGT 4560
ACGCTGCTAC ATGTCGGCAC TCCACGCAA CAAAAGAAG CCCAACCGAA AACGCACGCG 4620
CCTTTCCAGG CTCACCACGG AAAAAAATAC CACGCGCCGC TCACGAGCAA ACCGTGACAA 4680
60 CAGCCAGCCA GATATGGCAA CGGAGGCACG GGCGCACAC AGCCACTGAA AACCGCAGCT 4740
GCTCTTCCGT CCGTCCGTCC CTCCGCCGT CCGCGCCACT CCACTCGCCT TGCCCCACTC 4800
65 CCACTCTTCT CTCCCCGCGC ACACCGAGTC GGCACCGGCT CATCACCCAT CACCTCGGCC 4860
TCGGCCACCG GCAAACCCCC CGATCCGCTT TTGCAGGCAG CGCACTAAAA CCCCGGGGAG 4920

- 100 -

CGCGCCCCGC GGCAGCAGCA GCACCGCAGT GGGAGAGAGA GGCTTCGCC CGGCCCCCAC 4980
 CGAGCGGGGC GATCCACCGT CCGTGCCTCC GCACCTCCTC CGCCTCCTCC CCTGTCCCAC 5040
 5 GCGCCCACAC CCATGGCGGC GACGGCGTC GG 5072

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1706 base pairs
 10 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
 15 (iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: triticum tauschii
 20 (F) TISSUE TYPE: Endosperm

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 1..1706
 25 (D) OTHER INFORMATION:/product= "partial cDNA for hexaploid wheat DBE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

| | | |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 30 | GCT GTG TCG AAG CTT GAC TAT TTG AAG GAG CTT GGA GTT AAT TGT ATT Ala Val Ser Lys Leu Asp Tyr Leu Lys Glu Leu Gly Val Asn Cys Ile 1 5 10 15 | 48 |
| 35 | GAA TTA ATG CCC TGC CAT GAG TTC AAC GAG CTG GAG TAC TCA ACC TCT Glu Leu Met Pro Cys His Glu Phe Asn Glu Leu Glu Tyr Ser Thr Ser 20 25 30 | 96 |
| 40 | TCT TCC AAG ATG AAC TTT TGG GGA TAT TCT ACC ATA AAC TTC TTT TCA Ser Ser Lys Met Asn Phe Trp Gly Tyr Ser Thr Ile Asn Phe Phe Ser 35 40 45 | 144 |
| 45 | CCA ATG ACG AGA TAC ACA TCA GGC GGG ATA AAA AAC TGT GGG CGT GAT Pro Met Thr Arg Tyr Thr Ser Gly Gly Ile Lys Asn Cys Gly Arg Asp 50 55 60 | 192 |
| 50 | GCC ATA AAT GAG TTC AAA ACT TTT GTA AGA GAG GCT CAC AAA CGG GGA Ala Ile Asn Glu Phe Lys Thr Phe Val Arg Glu Ala His Lys Arg Gly 65 70 75 80 | 240 |
| 55 | ATT GAG GTG ATC CTG GAT GTT GTC TTC AAC CAT ACA GCT GAG GGT AAT Ile Glu Val Ile Leu Asp Val Val Phe Asn His Thr Ala Glu Gly Asn 85 90 95 | 288 |
| 60 | GAG AAT GGT CCA ATA TTA TCA TTT AGG GGG GTC GAT AAT ACT ACA TAC Glu Asn Gly Pro Ile Leu Ser Phe Arg Gly Val Asp Asn Thr Thr Tyr 100 105 110 | 336 |
| | TAT ATG CTT GCA CCC AAG GGA GAG TTT TAT AAC TAT TCT GGC TGT GGG Tyr Met Leu Ala Pro Lys Gly Glu Phe Tyr Asn Tyr Ser Gly Cys Gly 115 120 125 | 384 |

- 101 -

| | | |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | AAT ACC TTC AAC TGT AAT CAT CCT GTG GTT CGT CAA TTC ATT GTA GAT Asn Thr Phe Asn Cys Asn His Pro Val Val Arg Gln Phe Ile Val Asp 130 135 140 | 432 |
| 5 | TGT TTA AGA TAC TGG GTG ATG GAA ATG CAT GTT GAT GGT TTT CGT TTT Cys Leu Arg Tyr Trp Val Met Glu Met His Val Asp Gly Phe Arg Phe 145 150 155 160 | 480 |
| 10 | GAT CTT GCA TCC ATA ATG ACC AGA GGT TCC AGT CTG TGG GAT CCA GTT Asp Leu Ala Ser Ile Met Thr Arg Gly Ser Ser Leu Trp Asp Pro Val 165 170 175 | 528 |
| 15 | AAC GTG TAT GGA GCT CCA ATA GAA GGT GAC ATG ATC ACA ACA GGG ACA Asn Val Tyr Gly Ala Pro Ile Glu Gly Asp Met Ile Thr Thr Gly Thr 180 185 190 | 576 |
| 20 | CCT CTT GTT ACT CCA CCA CTT ATT GAC ATG ATC AGC AAT GAC CCA ATT Pro Leu Val Thr Pro Pro Leu Ile Asp Met Ile Ser Asn Asp Pro Ile 195 200 205 | 624 |
| 25 | CTT GGA GGC GTC AAG CTC ATT GCT GAA GCA TGG GAT GCA GGA GGC CTC Leu Gly Val Lys Leu Ile Ala Glu Ala Trp Asp Ala Gly Gly Leu 210 215 220 | 672 |
| 30 | TAT CAA GTA GGT CAA TTC CCT CAC TGG AAT GTT TGG TCT GAG TGG AAT Tyr Gln Val Gly Gln Phe Pro His Trp Asn Val Trp Ser Glu Trp Asn 225 230 235 240 | 720 |
| 35 | GGG AAG TAC CGG GAC ATT GTG CGC CAA TTC ATT AAA GGC ACT GAT GGA Gly Lys Tyr Arg Asp Ile Val Arg Gln Phe Ile Lys Gly Thr Asp Gly 245 250 255 | 768 |
| 40 | TTT GCT GGT GGT TTT GCC GAA TGT CTT TGT GGA AGT CCA CAC CTA TAC Phe Ala Gly Gly Phe Ala Glu Cys Leu Cys Gly Ser Pro His Leu Tyr 260 265 270 | 816 |
| 45 | CAG GCA GGA GGA AGG AAA CCT TGG CAC AGT ATC AAC TTT GTA TGT GCA Gln Ala Gly Gly Arg Lys Pro Trp His Ser Ile Asn Phe Val Cys Ala 275 280 285 | 864 |
| 50 | CAT GAT GGA TTT ACA CTG GGT GAT TTG GTA ACA TAT AAT AAC AAG TAC His Asp Gly Phe Thr Leu Gly Asp Leu Val Thr Tyr Asn Asn Lys Tyr 290 295 300 | 912 |
| 55 | AAT TTA CCA AAT GGG GAG AAC AAT AGA GAT GGA GAA AAT CAC AAT CTT Asn Leu Pro Asn Gly Glu Asn Asn Arg Asp Gly Glu Asn His Asn Leu 305 310 315 320 | 960 |
| 60 | AGC TGG AAT TGT GGG GAG GAA GGA GAA TTC GCA AGA TTG TCT GTC AAA Ser Trp Asn Cys Gly Glu Glu Gly Glu Phe Ala Arg Leu Ser Val Lys 325 330 335 | 1008 |
| | AGA TTG AGG AAG AGG CAG ATG CGC AAT TTC TTT GTT TGT CTC ATG GTT Arg Leu Arg Lys Arg Gln Met Arg Asn Phe Phe Val Cys Leu Met Val 340 345 350 | 1056 |
| | TCT CAA GGA GTT CCA ATG TTT TAC ATG GGC GAT GAA TAT GGC CAC ACA Ser Gln Gly Val Pro Met Phe Tyr Met Gly Asp Glu Tyr Gly His Thr 355 360 365 | 1104 |
| | AAA GGG GGC AAC AAC AAT ACA TAC TGC CAT GAT TCT TAT GTC AAT TAT Lys Gly Gly Asn Asn Asn Thr Tyr Cys His Asp Ser Tyr Val Asn Tyr 370 375 380 | 1152 |

- 102 -

| | | |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | TTT CGC TGG GAT AAA AAA GAA CAA TAC TCT GAC TTG CAC AGA TTC TGC Phe Arg Trp Asp Lys Lys Glu Gln Tyr Ser Asp Leu His Arg Phe Cys 385 390 395 400 | 1200 |
| 5 | TGC CTC ATG ACC AAA TTC CGC AAG GAG TGC GAG GGT CTT GGC CTT GAG Cys Leu Met Thr Lys Phe Arg Lys Glu Cys Glu Gly Leu Gly Leu Glu 405 410 415 | 1248 |
| 10 | GAC TTT CCA ACG GCC GAA CGG CTG CAG TGG CAT GGT CAT CAG CCT GGG Asp Phe Pro Thr Ala Glu Arg Leu Gln Trp His Gly His Gln Pro Gly 420 425 430 | 1296 |
| 15 | AAG CCT GAT TGG TCT GAG AAT AGC CGA TTC GTT GCC TTT TCC ATG AAA Lys Pro Asp Trp Ser Glu Asn Ser Arg Phe Val Ala Phe Ser Met Lys 435 440 445 | 1344 |
| 20 | GAT GAA AGA CAG GGC GAG ATC TAT GTG GCC TTC AAC ACC AGC CAC TTA Asp Glu Arg Gln Gly Glu Ile Tyr Val Ala Phe Asn Thr Ser His Leu 450 455 460 | 1392 |
| 25 | CCG GCC GTT GTT GAG CTC CCA GAG CGC GCA GGG CGC CGG TGG GAA CCG Pro Ala Val Val Glu Leu Pro Glu Arg Ala Gly Arg Arg Trp Glu Pro 465 470 475 480 | 1440 |
| 30 | GTG GTG GAC ACA GGC AAG CCA GCA CCA TAT GAC TTC CTC ACC GAC GAC Val Val Asp Thr Gly Lys Pro Ala Pro Tyr Asp Phe Leu Thr Asp Asp 485 490 495 | 1488 |
| 35 | TTA CCT GAT CGC GCT CTC ACC ATA CAC CAG TTC TCT CAT TTC CTC AAC Leu Pro Asp Arg Ala Leu Thr Ile His Gln Phe Ser His Phe Leu Asn 500 505 510 | 1536 |
| 40 | TCC AAC CTC TAC CCC ATG CTC AGC TAC TCA TCG GTC ATC CTA GTA TTG Ser Asn Leu Tyr Pro Met Leu Ser Tyr Ser Val Ile Leu Val Leu 515 520 525 | 1584 |
| 45 | CGC CCT GAT GTT TGA GAG ACA AAT ATA TAC AGT AAA TAA TAT GTC TAT Arg Pro Asp Val * Glu Thr Asn Ile Tyr Ser Lys * Tyr Val Tyr 530 535 540 | 1632 |
| 50 | ATG TAG TCC TTT GGC GTA TTA TCA GTG TGC ACA ATT GCT CTA TTG CCA Met * Ser Phe Gly Val Leu Ser Val Cys Thr Ile Ala Leu Leu Pro 545 550 555 560 | 1680 |
| 55 | GTG ATC TAT TCG ATA GCG GCC GCG AA Val Ile Tyr Ser Ile Ala Ala Ala 565 | 1706 |
| 60 | (2) INFORMATION FOR SEQ ID NO: 17: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 9289 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: DNA (genomic) (iii) HYPOTHETICAL: NO (vi) ORIGINAL SOURCE: | |

- 103 -

(A) ORGANISM: *triticum tauschii*
(F) TISSUE TYPE: Endosperm

(ix) FEATURE:

5 (A) NAME/KEY: CDS
 (B) LOCATION: 1..9289
 (D) OTHER INFORMATION:/product= "genomic sequence of DBE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

| | | |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 10 | CGG GAC CGT CCC TTG GCA ACT TGG GTT ACG TTG GGA CCT GAC GCT TCG Arg Asp Arg Pro Leu Ala Thr Trp Val Thr Leu Gly Pro Asp Ala Ser 570 575 580 | 48 |
| 15 | CTT ATC CGG TGT GCC CTG AGA CGA GAT ATG TGC AGC TCC TAT CGG ATT Leu Ile Arg Cys Ala Leu Arg Arg Asp Met Cys Ser Ser Tyr Arg Ile 585 590 595 600 | 96 |
| 20 | TGT CGG CAC ATT CGG CGG CTT TGC TGG TCT TGT TTT ACC ATT GTC GAA Cys Arg His Ile Arg Arg Leu Cys Trp Ser Cys Phe Thr Ile Val Glu 605 610 615 | 144 |
| 25 | ATG TCT TAT AAA CCG GGA TTC CGA GAC TGA TCG GGT CTT CCC GGG AGA Met Ser Tyr Lys Pro Gly Phe Arg Asp * Ser Gly Leu Pro Gly Arg 620 625 630 | 192 |
| 30 | AGG TTT ATC CTT CGT TGA CCG TGA GAG CTT ATA ATG GGC TAA GTT GGG Arg Phe Ile Leu Arg * Pro * Glu Leu Ile Met Gly * Val Gly 635 640 645 | 240 |
| 35 | ACA CCC CTG CAG GGT ATT ATC TTT CGA AAG CCG TGC CCG CGG TTA TGA Thr Pro Leu Gln Gly Ile Ile Phe Arg Lys Pro Cys Pro Arg Leu *650 655 660 | 288 |
| 40 | GCG AGA TGG GAA TTT GTT AAT GTC CGA TTG TAG AGA ACC TGT CAC TTG Gly Arg Trp Glu Phe Val Asn Val Arg Leu * Arg Thr Cys His Leu 665 670 675 680 | 336 |
| 45 | ACT TAA TTT AAA ATT CAT CAA CCG TGT GTG TAG CCG TGA TGG TCT CTT Thr * Phe Lys Ile His Gln Pro Cys Val * Pro * Trp Ser Leu 685 690 695 | 384 |
| 50 | TTC GGC GGA GTC CGG GAA GTG AAC ACG GTT TGA GTT ATG CAT GAA CGT Phe Gly Gly Val Arg Glu Val Asn Thr Val * Val Met His Glu Arg 700 705 710 | 432 |
| 55 | AAG TAG TTT CAG GAT CAC TCC TTG ATC ACT TCT AGC TCC GCG ACC GTT Lys * Phe Gln Asp His Ser Leu Ile Thr Ser Ser Ser Ala Thr Val 715 720 725 | 480 |
| 60 | GCG TTG TTT CTC TTC TCG CTC TCA TTT GCG TAT GTT AGC CAC CAT ATA Ala Leu Phe Leu Phe Ser Leu Ser Phe Ala Tyr Val Ser His His Ile 730 735 740 | 528 |
| 65 | TGC TTA GTG TCT GCT GCA GCT CCA CCT CAT TAC CCC TTC CTT TCC TAT Cys Leu Val Ser Ala Ala Ala Pro Pro His Tyr Pro Phe Leu Ser Tyr 745 750 755 760 | 576 |
| 70 | AAG CTT AAA TAG TCT TGA TCT CGC GGG TGT GAG ATT GCT GAG TCC TCG Lys Leu Lys * Ser * Ser Arg Gly Cys Glu Ile Ala Glu Ser Ser 765 770 775 | 624 |

- 104 -

| | | |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | TGA CTT ACA GAT TCT ACC AAA ACA GTT GCA GGT GTC GAC GAT GCC AGT * Leu Thr Asp Ser Thr Lys Thr Val Ala Gly Val Asp Asp Ala Ser 780 785 790 | 672 |
| 5 | GCA GGT GAC GCA ACC GAG CTC AAG TGG GAG TTC GAC GAG GAA CGT GGT Ala Gly Asp Ala Thr Glu Leu Lys Trp Glu Phe Asp Glu Glu Arg Gly 795 800 805 | 720 |
| 10 | CGT TAC TAT GTT TCT TTT CCT GAT GAT CAG TAG TGG AGC CCA GTT GGG Arg Tyr Tyr Val Ser Phe Pro Asp Asp Gln * Trp Ser Pro Val Gly 810 815 820 | 768 |
| 15 | ACG ATC GGG GAT CTA GCA TTT GGG GTT ATC TTA ATT TCT TTT AGA TTT Thr Ile Gly Asp Leu Ala Phe Gly Val Ile Leu Ile Ser Phe Arg Phe 825 830 835 840 | 816 |
| 20 | GAC CGT AAT CGG TCT ATG TGT GGA TTT TGG ATG ATG TAT GAA TTA TTT Asp Arg Asn Arg Ser Met Cys Gly Phe Trp Met Met Tyr Glu Leu Phe 845 850 855 | 864 |
| | ATG TAT TGT GTG AAG TGG CGA TTG TAA GCC AAC TCT CGT TAT CCC ATT Met Tyr Cys Val Lys Trp Arg Leu * Ala Asn Ser Arg Tyr Pro Ile 860 865 870 | 912 |
| 25 | CTT GTT CAT TAC ATG GGA TTG TGT GAA GAT GAC CCT TCT TGC GAC AAA Leu Val His Tyr Met Gly Leu Cys Glu Asp Asp Pro Ser Cys Asp Lys 875 880 885 | 960 |
| 30 | ACC ACA ATG CGG TTA TGC CTC TAA GTC GTG CCT CGA CAC GTG GGA GAT Thr Thr Met Arg Leu Cys Leu * Val Val Pro Arg His Val Gly Asp 890 895 900 | 1008 |
| 35 | ATA GCC GCA TCG TGG GCG TTA CAC GCA AGT CTT CAT AGC AAC CAA AAC Ile Ala Ala Ser Trp Ala Leu His Ala Ser Leu His Ser Asn Gln Asn 905 910 915 920 | 1056 |
| 40 | TCC TCT CCG CAT TAC AAG CCA CCA ATC GCA GCC ACC ATG ACT TTC TTC Ser Ser Pro His Tyr Lys Pro Pro Ile Ala Ala Thr Met Thr Phe Phe 925 930 935 | 1104 |
| | ACC ACT GTC AAT GCC ATG AAA ATC TAT ATG TAG ACA TGT CCC ATT GCA Thr Thr Val Asn Ala Met Lys Ile Tyr Met * Thr Cys Pro Ile Ala 940 945 950 | 1152 |
| 45 | TCG GCA AGA AAG CGA AGC TTC ACG GCA CAC CTT CAT GAA GCC TCT CTG Ser Ala Arg Lys Arg Ser Phe Thr Ala His Leu His Glu Ala Ser Leu 955 960 965 | 1200 |
| 50 | GCC GAA GAC AAG GAT GCG CCC GAC CGG ATC AAT TCC TAT CTA GAT ACC Ala Glu Asp Lys Asp Ala Pro Asp Arg Ile Asn Ser Tyr Leu Asp Thr 970 975 980 | 1248 |
| 55 | TAG TGG AGC CAT GCG CCA ATA GCG GAG ATC TCC GAG AGG AAG ACC GGA * Trp Ser His Ala Pro Ile Ala Glu Ile Ser Glu Arg Lys Thr Gly 985 990 995 1000 | 1296 |
| 60 | ACT CGT CGG ACG TCG GCG TCC AAA TCG AGG AGG CCG GCA TGA AGC ACA Thr Arg Arg Thr Ser Ala Ser Lys Ser Arg Arg Pro Ala * Ser Thr 1005 1010 1015 | 1344 |
| | TCG AGG ATG GTG ATC CCC ATA CGG GTA GAT CGG GTC GGC CGC CAT CTC Ser Arg Met Val Ile Pro Ile Arg Val Asp Arg Val Gly Arg His Leu 1020 1025 1030 | 1392 |

- 105 -

| | | |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | ACA CCG AGA TTA GGA TGC TTA AAA CGG TTT TGG CAC TAG CAT TAT Thr Pro Arg Leu Gly Cys Leu Lys Arg Phe Phe Trp His * His Tyr 1035 1040 1045 | 1440 |
| 5 | TTT GCA TCA TCC GTT GGA GAG AAC ATG AGA GAG CCC CAT TTC TTC CAC Phe Ala Ser Ser Val Gly Glu Asn Met Arg Glu Pro His Phe Phe His 1050 1055 1060 | 1488 |
| 10 | GGT TCT ACC TAT GGG ATC TTG TTC TGC TTG CAA CCG GGC CTC ACG GAA Gly Ser Thr Tyr Gly Ile Leu Phe Cys Leu Gln Pro Gly Leu Thr Glu 1065 1070 1075 1080 | 1536 |
| 15 | AAC CCG CGC CAG CGG ACC CAC CCC ATG CTA GCA GGG CAC GGC ACC CGC Asn Pro Arg Gln Arg Thr His Pro Met Leu Ala Gly His Gly Thr Arg 1085 1090 1095 | 1584 |
| 20 | AGC GGC CGG TCC AAA TGG ACG GTG AGA ACC GCA ACG CGA CAC GCC CGG Ser Gly Arg Ser Lys Trp Thr Val Arg Thr Ala Thr Arg His Ala Arg 1100 1105 1110 | 1632 |
| 25 | CAC TGT CAG CAA AGC GAG AGC GCG CGC ACG GCA CAC GCA CGC TCG GAC His Cys Gln Gln Ser Glu Ser Ala Arg Thr Ala His Ala Arg Ser Asp 1115 1120 1125 | 1680 |
| 30 | GAA CGG ACG GTG CGA TCG ATC CCT CCC CCC TCG CTC AAC CAC AGT AGT Glu Arg Thr Val Arg Ser Ile Pro Pro Ser Leu Asn His Ser Ser 1130 1135 1140 | 1728 |
| 35 | ACC CTG CCA CAC TAT CAC GCA CGC ACT CGA GTC ACA CCT CCC ACG AAG Thr Leu Pro His Tyr His Ala Arg Thr Arg Val Thr Pro Pro Thr Lys 1145 1150 1155 1160 | 1776 |
| 40 | AAC CAA CAG GAG GCG CGG ATC CCA CCG ATA AAT AAC CCC GCC TCG CCG Asn Gln Gln Glu Ala Arg Ile Pro Pro Ile Asn Asn Pro Ala Ser Pro 1165 1170 1175 | 1824 |
| 45 | CTC CTC CCC AAA ATC AAT CAC CGA TCG CTC GGG GTT CCC GGC ATG ACG Leu Leu Pro Lys Ile Asn His Arg Ser Leu Gly Val Pro Gly Met Thr 1180 1185 1190 | 1872 |
| 50 | ATG ATG GCC ATG GCC AAG GCG CCC TGC CTC TGC GCG CGC CCG TCC CTC Met Met Ala Met Ala Lys Ala Pro Cys Leu Cys Ala Arg Pro Ser Leu 1195 1200 1205 | 1920 |
| 55 | GCC GCG CGC GCG AGG CGG CCG GGG CCG GGG CCG GCG CCG CGC CTG CGA Ala Ala Arg Ala Arg Arg Pro Gly Pro Gly Pro Ala Pro Arg Leu Arg 1210 1215 1220 | 1968 |
| 60 | CGG TGG CGA CCC AAT GCG ACG GCG GGG AAG GGG GTC GGC GAG GTG TGC Arg Trp Arg Pro Asn Ala Thr Ala Gly Lys Gly Val Gly Glu Val Cys 1225 1230 1235 1240 | 2016 |
| | GCC GCG GTT GTC GAG GCG GCG ACG AAG GCC GAG GAT GAG GAC GAC GAC Ala Ala Val Val Glu Ala Ala Thr Lys Ala Glu Asp Glu Asp Asp Asp 1245 1250 1255 | 2064 |
| | GAG GAG GAG GCG GTG GCG GAG GAC AGG TAC GCG CTC GGC GGC GCG TGC Glu Glu Glu Ala Val Ala Glu Asp Arg Tyr Ala Leu Gly Gly Ala Cys 1260 1265 1270 | 2112 |
| | AGG GTG CTC GCC GGA ATG CCC GCG CCG CTG GGC GCC ACC GCG CTC GCC Arg Val Leu Ala Gly Met Pro Ala Pro Leu Gly Ala Thr Ala Leu Ala 1275 1280 1285 | 2160 |

- 106 -

| | | |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | GGC GGG GTC AAT TTC GCC GTC TAC TCC GGT GGA GCC ACC GCC GCG GCG Gly Gly Val Asn Phe Ala Val Tyr Ser Gly Gly Ala Thr Ala Ala Ala 1290 1295 1300 | 2208 |
| 5 | CTC TGC CTC TTC ACG CCA GAA GAT CTC AAG GCG GTG GGG TTG CCT CCC Leu Cys Leu Phe Thr Pro Glu Asp Leu Lys Ala Val Gly Leu Pro Pro 1305 1310 1315 1320 | 2256 |
| 10 | GAG TAG AGT TCA TCA GCT TTG CGT GCG CCG CGC GCC CCC TTT TCT GGC Glu * Ser Ser Ser Ala Leu Arg Ala Pro Arg Ala Pro Phe Ser Gly 1325 1330 1335 | 2304 |
| 15 | CTG CGA TTT AAG TTT TGT ACT GGG GGA AAT GCT GCA GGA TAG GGT GAC Leu Arg Phe Lys Phe Cys Thr Gly Gly Asn Ala Ala Gly * Gly Asp 1340 1345 1350 | 2352 |
| 20 | GGA GGA GGT TTC CCT TGA CCC CCT GAT GAA TCG GAC TGG GAA CGT GTG Gly Gly Phe Pro * Pro Pro Asp Glu Ser Asp Trp Glu Arg Val 1355 1360 1365 | 2400 |
| 25 | GCA TGT CTT CAT TGA AGG CGA GCT GCA CGA CAT GCT TTA CGG GTA CAG Ala Cys Leu His * Arg Arg Ala Ala Arg His Ala Leu Arg Val Gln 1370 1375 1380 | 2448 |
| 30 | GTT CGA CGG CAC CTT TGC TCC TCA CTG CGG GCA CTA CCT TGA TAT TTC Val Arg Arg His Leu Cys Ser Ser Leu Arg Ala Leu Pro * Tyr Phe 1385 1390 1395 1400 | 2496 |
| 35 | CAA TGT CGT GGT GGA TCC TTA TGC TAA GGT GAT CAT ACT TTA GCT TTA Gln Cys Arg Gly Gly Ser Leu Cys * Gly Asp His Thr Leu Ala Leu 1405 1410 1415 | 2544 |
| 40 | CCT GCA TCT TGG TAT TTA CAG TAG AAA TTG TTA CGT GGA CCC TTA TTT Pro Ala Ser Trp Tyr Leu Gln * Lys Leu Leu Arg Gly Pro Leu Phe 1420 1425 1430 | 2592 |
| 45 | GTT GCC TTT TGT GTT GCT CTA GGC AGT GAT AAG CCG AGG GGA GTA TGG Val Ala Phe Cys Val Ala Leu Gly Ser Asp Lys Pro Arg Gly Val Trp 1435 1440 1445 | 2640 |
| 50 | CGT TCC GGC GCG TGG TAA CAA TTG CTG GCC TCA GAT GGC TGG CAT GAT Arg Ser Gly Ala Trp * Gln Leu Leu Ala Ser Asp Gly Trp His Asp 1450 1455 1460 | 2688 |
| 55 | CCC TCT TCC ATA TAG CAC GGT ATG CCT GAT TGC TGA AAA TAT TGG CTG Pro Ser Ser Ile * His Gly Met Pro Asp Cys * Lys Tyr Trp Leu 1465 1470 1475 1480 | 2736 |
| 60 | CAT TTG TTT CTC TCT TTT CAT ATT TTT CTC CTG TCT TTC ACT TGT His Leu Phe Leu Ser Phe Ser His Ile Phe Leu Leu Ser Phe Thr Cys 1485 1490 1495 | 2784 |
| | ACT ACA TTG CCT CAG ACA GTC ATG ATC AAA GAG AGC AGT GTC ATT AGA Thr Thr Leu Pro Gln Thr Val Met Ile Lys Glu Ser Ser Val Ile Arg 1500 1505 1510 | 2832 |
| | CAT TTG TAG TTG TCT GCT GAC TTT GAC CAA AAC TTG TAA TTT ACT GTT His Leu * Leu Ser Ala Asp Phe Asp Gln Asn Leu * Phe Thr Val 1515 1520 1525 | 2880 |
| | GTT AAA GGT CCT TGA ATC ATA TTT TTT TAT AAT ATT ATG TTT GCA AGT Val Lys Gly Pro * Ile Ile Phe Phe Tyr Asn Ile Met Phe Ala Ser 1530 1535 1540 | 2928 |

- 107 -

| | | |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | GGA AGT AAA GTG AAA TTG CAT CTA GTA TTT GTT GTT GCT GTC TTA GTC Gly Ser Lys Val Lys Leu His Leu Val Phe Val Val Ala Val Leu Val 1545 1550 1555 1560 | 2976 |
| 5 | GTT TAA TTG GAC ATG CAG TAA AAA GGT TTG CAT CTG CAG TTT GAT TGG Val * Leu Asp Met Gln * Lys Gly Leu His Leu Gln Phe Asp Trp 1565 1570 1575 | 3024 |
| 10 | GAA GGC GAC CTA CCT CTA AGA TAT CCT CAA AAG GAC CTG GTA ATA TAT Glu Gly Asp Leu Pro Leu Arg Tyr Pro Gln Lys Asp Leu Val Ile Tyr 1580 1585 1590 | 3072 |
| 15 | GAG ATG CAC TTG CGT GGA TTC ACG AAG CAT GAT TCA AGC AAT GTA GAA Glu Met His Leu Arg Gly Phe Thr Lys His Asp Ser Ser Asn Val Glu 1595 1600 1605 | 3120 |
| 20 | CAT CCG GGT ACT TTC ATT GGA GCT GTG TCG AAG CTT GAC TAT TTG AAG His Pro Gly Thr Phe Ile Gly Ala Val Ser Lys Leu Asp Tyr Leu Lys 1610 1615 1620 | 3168 |
| 25 | GTA CAG CTG TAC TTG CTG ACT ACA TAG GAT AAT TTT TAA AGA AAG CTA Val Gln Leu Tyr Leu Thr Thr * Asp Asn Phe * Arg Lys Leu 1625 1630 1635 1640 | 3216 |
| 30 | CAT ATT AGC CAG AAT TTG GGT TAT TAC AAA AAC TAC TGC ATA CTA TAG His Ile Ser Gln Asn Leu Gly Tyr Tyr Lys Asn Tyr Cys Ile Leu *1645 1650 1655 | 3264 |
| 35 | CAG TTA CAT GCT CAT TAT CGA GGA GAT GCT CAC ACG CAT CTT ATT TGG Gln Leu His Ala His Tyr Arg Gly Asp Ala His Thr His Leu Ile Trp 1660 1665 1670 | 3312 |
| 40 | ATT TAA TAC CCA ATT CTG TTT TGA TAT TGG ACT GTT CCC TCT ACA GGA Ile * Tyr Pro Ile Leu Phe * Tyr Trp Thr Val Pro Ser Thr Gly 1675 1680 1685 | 3360 |
| 45 | GCT TGG AGT TAA TTG TAT TGA ATT AAT GCC CTG CCA TGA GTT CAA CGA Ala Trp Ser * Leu Tyr * Ile Asn Ala Leu Pro * Val Gln Arg 1690 1695 1700 | 3408 |
| 50 | GCT GGA GTA CTC AAC CTC TTC TTC CAA GTA AGG ACA TGA ATT TAG TAT Ala Gly Val Leu Asn Leu Phe Phe Gln Val Arg Thr * Ile * Tyr 1705 1710 1715 1720 | 3456 |
| 55 | TAG CCT GCC AGC ACT GTT TGA GTG AGA GTT CAT ACA CAT TTT GTG CCT * Pro Ala Ser Thr Val * Val Arg Val His Thr His Phe Val Pro 1725 1730 1735 | 3504 |
| 60 | GCA TAA CTG ATA TTT GTT CAA ACT ATT TTT TTT AGC AGT CAC TCA ACA Ala * Leu Ile Phe Val Gln Thr Ile Phe Phe Ser Ser His Ser Thr 1740 1745 1750 | 3552 |
| | GTT TTA CAT ATA TAT ATA ATA TAG ACT ATT CGT CAC CCT GGG TGA GGA Val Leu His Ile Tyr Ile Ile * Thr Ile Arg His Pro Gly * Gly 1755 1760 1765 | 3600 |
| | ATA GTT ATT CTT CAC CCA CCT CTA TTT TAA CAT CTA TGC ACC GTA ATT Ile Val Ile Leu His Pro Pro Leu Phe * His Leu Cys Thr Val Ile 1770 1775 1780 | 3648 |
| | TTA CGT TTC GTA AAT TTG TCT TAT TTT AGA GAT AAA AAG AGA ACG TAA Leu Arg Phe Val Asn Leu Ser Tyr Phe Arg Asp Lys Lys Arg Thr * 1785 1790 1795 1800 | 3696 |

- 108 -

| | | |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | GAA AAC CTA TAA TCG TCG TAA AAA AAA ATA TGT TAC GTA AAA TTA CAA Glu Asn Leu * Ser Ser * Lys Lys Ile Cys Tyr Val Lys Leu Gln 1805 1810 1815 | 3744 |
| 5 | ATG TAA AAA CAT AGT GTA AAA TGT ACA TAA AAT ACA TTT TTT GAC CTA Met * Lys His Ser Val Lys Cys Thr * Asn Thr Phe Phe Asp Leu 1820 1825 1830 | 3792 |
| 10 | TAT TTT TTT TGT TAA TGC CAA ATT TTA TAC AGT AAA TCA ATA TGA ATG Tyr Phe Phe Cys * Cys Gln Ile Leu Tyr Ser Lys Ser Ile * Met 1835 1840 1845 | 3840 |
| 15 | TAA CTA TTT GTA TTT CAA ATG TAA TTT ATT TAT GAA ATG GTC GTA AGA * Leu Phe Val Phe Gln Met * Phe Ile Tyr Glu Met Val Val Arg 1850 1855 1860 | 3888 |
| 20 | TTA CCT CGG GTG AAG AAT AAC TTA TTC TGC ACC CTG GGT GAT GAA TAG Leu Pro Arg Val Lys Asn Asn Leu Phe Cys Thr Leu Gly Asp Glu * 1865 1870 1875 1880 | 3936 |
| 25 | TAA CAC TAT ATA TAT ATA TAT ATA TAT ATA TAT ATA CCG GCT * His Tyr Ile Tyr Ile Tyr Ile Tyr Ile Tyr Ile Tyr Ile Pro Ala 1885 1890 1895 | 3984 |
| 30 | GCT GCT AAT GAT GTT AAT ATT TCG CAA GTA CCT AAG CTG GAT TTT TCT Ala Ala Asn Asp Val Asn Ile Ser Gln Val Pro Lys Leu Asp Phe Ser 1900 1905 1910 | 4032 |
| 35 | CCA TGA GAC ATC AAT CCA TAA TTG AAA TTG GTC ACG ACA GTT GAA TAG Pro * Asp Ile Asn Pro * Leu Lys Leu Val Thr Thr Val Glu * 1915 1920 1925 | 4080 |
| 40 | TTG ATA GCT GAA AAT GAA ATC CAG CAT GCT ACT GTC TTG CCA TCT CCA Leu Ile Ala Glu Asn Glu Ile Gln His Ala Thr Val Leu Pro Ser Pro 1930 1935 1940 | 4128 |
| 45 | GAC TTG CTA ACA TGA ATT TTG TCT GCC TAC CTG TCA TTT GTA CCA ACG Asp Leu Leu Thr * Ile Leu Ser Ala Tyr Leu Ser Phe Val Pro Thr 1945 1950 1955 1960 | 4176 |
| 50 | TTC CCA ATT GCC CTC TCA TTA TTC GTG TGT ACC ATG CAT ATG TGT TTT Phe Pro Ile Ala Leu Ser Leu Phe Val Cys Thr Met His Met Cys Phe 1965 1970 1975 | 4224 |
| 55 | AAC ATG ATT ATT GTT GGC TAT ATT TCT CTT TGG AAA CAT GAC TAA TTT Asn Met Ile Ile Val Gly Tyr Ile Ser Leu Trp Lys His Asp * Phe 1980 1985 1990 | 4272 |
| 60 | ATC ACC CGT TTT GTA TAA ACT GCT TGT TTT CAT ATC AGG ATG AAC TTT Ile Thr Arg Phe Val * Thr Ala Cys Phe His Ile Arg Met Asn Phe 1995 2000 2005 | 4320 |
| | TGG GGA TAT TCT ACC ATA AAC TTC TTT TCA CCA ATG ACG AGA TAC ACA Trp Gly Tyr Ser Thr Ile Asn Phe Phe Ser Pro Met Thr Arg Tyr Thr 2010 2015 2020 | 4368 |
| | TCA GGC GGG ATA AAA AAC TGT GGG CGT GAT GCC ATA AAT GAG TTC AAA Ser Gly Gly Ile Lys Asn Cys Gly Arg Asp Ala Ile Asn Glu Phe Lys 2025 2030 2035 2040 | 4416 |
| | ACT TTT GTA AGA GAG GCT CAC AAA CGG GGA ATT GAG GTA AGC AAG TCG Thr Phe Val Arg Glu Ala His Lys Arg Gly Ile Glu Val Ser Lys Ser 2045 2050 2055 | 4464 |

- 109 -

| | | |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | TAC GAG TTA GTT GCT CCT TTT GAA CTT ATC AAT TTG ATG CGA AGA CAT Tyr Glu Leu Val Ala Pro Phe Glu Leu Ile Asn Leu Met Arg Arg His 2060 2065 2070 | 4512 |
| 5 | GTT ACT GCT AGG TGA TCC TGG ATG TTG TCT TCA ACC ATA CAG CTG AGG Val Thr Ala Arg * Ser Trp Met Leu Ser Ser Thr Ile Gln Leu Arg 2075 2080 2085 | 4560 |
| 10 | GTA ATG AGA ATG GTC CAA TAT TAT CAT TTA GGG GGG TCG ATA ATA CTA Val Met Arg Met Val Gln Tyr Tyr His Leu Gly Gly Ser Ile Ile Leu 2090 2095 2100 | 4608 |
| 15 | CAT ACT ATA TGC TTG CAC CCA AGG TGA CAG ATC TTT CTT GCT GCG TAA His Thr Ile Cys Leu His Pro Arg * Gln Ile Phe Leu Ala Ala * 2105 2110 2115 2120 | 4656 |
| 20 | TTG TTC TTT CAT AGA TGT ATA GAG CAT AGA TGT GTT ATG TAG TAG TTC Leu Phe His Arg Cys Ile Glu His Arg Cys Val Met * * Phe 2125 2130 2135 | 4704 |
| 25 | TTT TTC AAG GGG ATT ATG TTC ATG CAG GGA GAG TTT TAT AAC TAT TCT Phe Phe Lys Gly Ile Met Phe Met Gln Gly Glu Phe Tyr Asn Tyr Ser 2140 2145 2150 | 4752 |
| 30 | GGC TGT GGG AAT ACC TTC AAC TGT AAT CAT CCT GTG GTT CGT CAA TTC Gly Cys Gly Asn Thr Phe Asn Cys Asn His Pro Val Val Arg Gln Phe 2155 2160 2165 | 4800 |
| 35 | ATT GTA GAT TGT TTA AGG TAC AGA TAT ACA TTT TAC TTC TAG AAC TAC Ile Val Asp Cys Leu Arg Tyr Arg Tyr Thr Phe Tyr Phe * Asn Tyr 2170 2175 2180 | 4848 |
| 40 | TTT TTC ATT TCT TTT GCT GCT TGT CAT TTT GAT ATG ATT AAT TTG CAA Phe Phe Ile Ser Phe Ala Ala Cys His Phe Asp Met Ile Asn Leu Gln 2185 2190 2195 2200 | 4896 |
| 45 | GCT TGT GGG GGT AAA TCT TTT GGT CAG CAT ATT GTA TCT TTA AAT GTC Ala Cys Gly Gly Lys Ser Phe Gly Gln His Ile Val Ser Leu Asn Val 2205 2210 2215 | 4944 |
| 50 | ACA AAT ACT AAT GTC CTG GTG CTT ATT GAT TTG GCA TCT TCA AAT TCT Thr Asn Thr Asn Val Leu Val Ile Asp Leu Ala Ser Ser Asn Ser 2220 2225 2230 | 4992 |
| 55 | TCT CCA ATG AAA AGG GAA AAA TCT ACT GTA TGT CTC GTC AAC TAA TTT Ser Pro Met Lys Arg Glu Lys Ser Thr Val Cys Leu Val Asn * Phe 2235 2240 2245 | 5040 |
| 60 | ACT TTT GTT TTG CAG ATA CTG GGT GAT GGA AAT GCA TGT TGA TGG TTT Thr Phe Val Leu Gln Ile Leu Gly Asp Gly Asn Ala Cys * Trp Phe 2250 2255 2260 | 5088 |
| | TCG TTT TGA TCT TGC ATC CAT AAT GAC CAG AGG TTC CAG GTA ATT TGT Ser Phe * Ser Cys Ile His Asn Asp Gln Arg Phe Gln Val Ile Cys 2265 2270 2275 2280 | 5136 |
| | ATT TAT TGT TTG TTT GCG TGT TGC CTT TTC AGA AGA TTC TTA AAA GAA Ile Tyr Cys Leu Phe Ala Cys Cys Leu Phe Arg Arg Phe Leu Lys Glu 2285 2290 2295 | 5184 |
| | TGT TTC TTT TAC AAG TCT GTG GGA TCC AGT TAA CGT GTA TGG AGC TCC Cys Phe Phe Tyr Lys Ser Val Gly Ser Ser * Arg Val Trp Ser Ser 2300 2305 2310 | 5232 |

- 110 -

| | | |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | AAT AGA AGG TGA CAT GAT CAC AAC AGG GAC ACC TCT TGT TAC TCC ACC Asn Arg Arg * His Asp His Asn Arg Asp Thr Ser Cys Tyr Ser Thr 2315 2320 2325 | 5280 |
| 5 | ACT TAT TGA CAT GAT CAG CAA TGA CCC AAT TCT TGG AGG CGT CAA GGT Thr Tyr * His Asp Gln Gln * Pro Asn Ser Trp Arg Arg Gln Gly 2330 2335 2340 | 5328 |
| 10 | ACT TGT TTC ATC CAA CAC CTG TTG TCT GTG TGC ATT CAA TTG TTT TAA Thr Cys Phe Ile Gln His Leu Leu Ser Val Cys Ile Gln Leu Phe * 2345 2350 2355 2360 | 5376 |
| 15 | TAT GGT AAT GAT CAA TTT CCC AAT GTT GAT AAG GAA AAA AAA TGC AAG Tyr Gly Asn Asp Gln Phe Pro Asn Val Asp Lys Glu Lys Lys Cys Lys 2365 2370 2375 | 5424 |
| 20 | TAG CTC TCT TTA TCT GCT TCT TGT GAG TTA TGC TAA ACA TGT AGA TAC * Leu Ser Leu Ser Ala Ser Cys Glu Leu Cys * Thr Cys Arg Tyr 2380 2385 2390 | 5472 |
| 25 | TAC TAT ATT TCA ACT GTA TAT ACT TGA CAT ATT ATT GCT TCC TTG GGA Tyr Tyr Ile Ser Thr Val Tyr Thr * His Ile Ile Ala Ser Leu Gly 2395 2400 2405 | 5520 |
| 30 | GGC TCT CTT ATT CCT TTC CCC CGT TGC AAT TAT AGC TCA TTG CTG AAG Gly Ser Leu Ile Pro Phe Pro Arg Cys Asn Tyr Ser Ser Leu Leu Lys 2410 2415 2420 | 5568 |
| 35 | CAT GGG ATG CAG GAG GCC TCT ATC AAG TAG GTC AAT TCC CTC ACT GGA His Gly Met Gln Glu Ala Ser Ile Lys * Val Asn Ser Leu Thr Gly 2425 2430 2435 2440 | 5616 |
| 40 | ATG TTT GGT CTG AGT GGA ATG GGA AGG TAA GGT ACC TGT TAA AAG TTT Met Phe Gly Leu Ser Gly Met Gly Arg * Gly Thr Cys * Lys Phe 2445 2450 2455 | 5664 |
| 45 | GAA TGG CAA ATA CTG ATA GAA ATA TAA CTT ATA TTT GCG ACA TAT ATA Glu Trp Gln Ile Leu Ile Glu Ile * Leu Ile Phe Ala Thr Tyr Ile 2460 2465 2470 | 5712 |
| 50 | GAT AAA GCA AAA TAA TAC GCA TTC CAC CTG AAC TTT AAA GGG GCA CGC Asp Lys Ala Lys * Tyr Ala Phe His Leu Asn Phe Lys Gly Ala Arg 2475 2480 2485 | 5760 |
| 55 | AGT GAA GTA CTA CTT CTC AAA TGT CTG AAT GAA CGC ACT AAC TCT TGT Ser Glu Val Leu Leu Leu Lys Cys Leu Asn Glu Arg Thr Asn Ser Cys 2505 2510 2515 2520 | 5808 |
| 60 | GAG TGT CAA CCG AGC AAG AAA TAT TTG AGT TTT CTG CAA GAA ATT GTT Glu Cys Gln Pro Ser Lys Lys Tyr Leu Ser Phe Leu Gln Glu Ile Val 2525 2530 2535 | 5856 |
| | CAT GTT GTG CTG TAT TAT ACT CCC TCC GTC CGA AAT TAT TTG TCG GAG His Val Val Leu Tyr Tyr Thr Pro Ser Val Arg Asn Tyr Leu Ser Glu 2540 2545 2550 | 5904 |
| | AAA TGG ATG TAT CTA GAC GTA TTT TAG TTC TAG ATA CAT CCA TTT TTA Lys Trp Met Tyr Leu Asp Val Phe * Phe * Ile His Pro Phe Leu 2555 2560 2565 | 5952 |
| | | 6000 |

- 111 -

| | | |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | TCC ATT TCT GCA ACA AGT AGT TCC GGA CGG AGG GAG TAT CAT TTA ACA Ser Ile Ser Ala Thr Ser Ser Ser Gly Arg Arg Glu Tyr His Leu Thr 2570 2575 2580 | 6048 |
| 5 | AAT ATA TGC ATG TTC GAA GTA AAT CCC CAC GAA TAA GCA TAT AAG ACG Asn Ile Cys Met Phe Glu Val Asn Pro His Glu * Ala Tyr Lys Thr 2585 2590 2595 2600 | 6096 |
| 10 | ATA TTG CTT TTT GAC TTG CAA CAC CTA AAC CTC ATT GTT TTC TCC TAG Ile Leu Leu Phe Asp Leu Gln His Leu Asn Leu Ile Val Phe Ser * 2605 2610 2615 | 6144 |
| 15 | GAT TTT GGG TGT TCG AAG CAA GCA GCT GGT GAT ATT TAA TTT ACC TTT Asp Phe Gly Cys Ser Lys Gln Ala Ala Gly Asp Ile * Phe Thr Phe 2620 2625 2630 | 6192 |
| 20 | GCC TTT ATT TGT AGC TTG ATT TGA GGG TGC GGC AAA GGT TTT AGC TTA Ala Phe Ile Cys Ser Leu Ile * Gly Cys Gly Lys Gly Phe Ser Leu 2635 2640 2645 | 6240 |
| 25 | GTA GTG TTT TGT AAA TTA TTA TAG TTT ATG TAT ATA CTC CTC ATT TGG Val Val Phe Cys Lys Leu Leu * Phe Met Tyr Ile Leu Leu Ile Trp 2650 2655 2660 | 6288 |
| 30 | GCA CTT CCG TAC TGG TCC CAT AGA AGA TAA AAA TGG AAT GAT GTC TGG Ala Leu Pro Tyr Trp Ser His Arg Arg * Lys Trp Asn Asp Val Trp 2665 2670 2675 2680 | 6336 |
| 35 | CCA ATA ATT GTT GAC AAC ACT GTT GCG CAT TTG ATT TTT ATC AGG GAA Pro Ile Ile Val Asp Asn Thr Val Ala His Leu Ile Phe Ile Arg Glu 2685 2690 2695 | 6384 |
| 40 | TGG AAA ATT GAA ATC GGT AAG AAA CAT TGC GAT ATT AAG CTT GTA TAT Trp Lys Ile Glu Ile Gly Lys Lys His Cys Asp Ile Lys Leu Val Tyr 2700 2705 2710 | 6432 |
| 45 | GCT AAT GCT GGT GGA TCT TTA AGA GGG AAC ATA TGA TCT CGT GTG CAT Ala Asn Ala Gly Gly Ser Leu Arg Gly Asn Ile * Ser Arg Val His 2715 2720 2725 | 6480 |
| 50 | CCA TCT TCA ACT AAA AAA ATA TGT TGC ACA TCT CCC ACG TCA CTT ACT Pro Ser Ser Thr Lys Lys Ile Cys Cys Thr Ser Pro Thr Ser Leu Thr 2730 2735 2740 | 6528 |
| 55 | AGC TAT TTC ATC CAA GTA CTA ACT TGT GTG GTT GTC TCC TCA GTA CCG Ser Tyr Phe Ile Gln Val Leu Thr Cys Val Val Val Ser Ser Val Pro 2745 2750 2755 2760 | 6576 |
| 60 | GGA CAT TGT GCG CCA ATT CAT TAA AGG CAC TGA TGG ATT TGC TGG TGG Gly His Cys Ala Pro Ile His * Arg His * Trp Ile Cys Trp Trp 2765 2770 2775 | 6624 |
| | TTT TGC CGA ATG TCT TTG TGG AAG TCC ACA CCT ATA CCA GGT AAG TTG Phe Cys Arg Met Ser Leu Trp Lys Ser Thr Pro Ile Pro Gly Lys Leu 2780 2785 2790 | 6672 |
| | TGG CAA TAC TTG GAA ATG GGT TGA GTG AAT GTC ACA TGG ATT TTT TAT Trp Gln Tyr Leu Glu Met Gly * Val Asn Val Thr Trp Ile Phe Tyr 2795 2800 2805 | 6720 |
| | ATA TAC CAC ATG ATG ATA CAC ATG TAA ATA TAT AAC GAT TAT AGT GTA Ile Tyr His Met Met Ile His Met * Ile Tyr Asn Asp Tyr Ser Val 2810 2815 2820 | 6768 |

- 112 -

| | | |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | TGC ATA TGC ATT TGG CTA AGA AGT ACT CCC TCC CTT AGT AAA AGT TAG Cys Ile Cys Ile Trp Leu Arg Ser Thr Pro Ser Leu Ser Lys Ser 2825 2830 2835 2840 | 6816 |
| 5 | TAC AAA GTT GAG TCA TCT ATT TTG GAA CGG AGG GAG TAT AAG TGT ATA Tyr Lys Val Glu Ser Ser Ile Leu Glu Arg Arg Glu Tyr Lys Cys Ile 2845 2850 2855 | 6864 |
| 10 | CAC TAG TGC AAT ATA TAG GTT TTA ACA CCC AAC TTG CCA ATG AAG GAA His * Cys Asn Ile * Val Leu Thr Pro Asn Leu Pro Met Lys Glu 2860 2865 2870 | 6912 |
| 15 | CAT AGG GCT TTC TAG TTA TCT TAT TTA TTT GTC TGG TGA ATA ATC CAC His Arg Ala Phe * Leu Ser Tyr Leu Phe Val Trp * Ile Ile His 2875 2880 2885 | 6960 |
| 20 | TGA AAA ATT CCA GCC ATG TCA TTT TTT AGG GGG GGA GAA GAA ACT ACA * Lys Ile Pro Ala Met Ser Phe Phe Arg Gly Gly Glu Glu Thr Thr 2890 2895 2900 | 7008 |
| 25 | TTG ATT TTT CCC CCT AAA AAA AGC CAT CTC AGA TTT CAT AGG TAA CTT Leu Ile Phe Pro Pro Lys Lys Ser His Leu Arg Phe His Arg * Leu 2905 2910 2915 2920 | 7056 |
| 30 | GCT TTT CTG TAA AGA AAT GAA AAC GAC TTC ATA CTT TCT GTC GAT TAT Ala Phe Leu * Arg Asn Glu Asn Asp Phe Ile Leu Ser Val Asp Tyr 2925 2930 2935 | 7104 |
| 35 | AAG TGT ATA CAC TAG TGC AAT ATA TAG GTT TTA ACA CCC AAC TTG CCA Lys Cys Ile His * Cys Asn Ile * Val Leu Thr Pro Asn Leu Pro 2940 2945 2950 | 7152 |
| 40 | ATG AAG GAA CAT AGG GCT TTC TAG TTA TCT TAT TTA TTT GCT GGT GAA Met Lys Glu His Arg Ala Phe * Leu Ser Tyr Leu Phe Ala Gly Glu 2955 2960 2965 | 7200 |
| 45 | TAA TCC ACT GAA AAA TTC CAG CCA TGT CAT TTT TTA GGG GGG AGA AGA * Ser Thr Glu Lys Phe Gln Pro Cys His Phe Leu Gly Gly Arg Arg 2970 2975 2980 | 7248 |
| 50 | AAC TAT ATT GAT TTT TCC CCC TAA AAA AAG CCA TCT CAG ATT CAT AGG Asn Tyr Ile Asp Phe Ser Pro * Lys Lys Pro Ser Gln Ile His Arg 2985 2990 2995 3000 | 7296 |
| 55 | AAC TTG CTT TTC TGT AAA GAA ATG AAA ACG ACT TCA TAC TTT CTG CGG Asn Leu Leu Phe Cys Lys Glu Met Lys Thr Thr Ser Tyr Phe Leu Arg 3005 3010 3015 | 7344 |
| 60 | CGC TTA CTT AGC TCG ATG GAT ATT TGT AAG ATG AAT GCC AAA TTA TTT Arg Leu Leu Ser Ser Met Asp Ile Cys Lys Met Asn Ala Lys Leu Phe 3020 3025 3030 | 7392 |
| 65 | GCC GGG ATT TGA TCG TTA TTC CAA ATT TCA TTT GGT TTC TCT AGC AAT Gly Gly Ile * Ser Leu Phe Gln Ile Ser Phe Gly Phe Ser Ser Asn 3035 3040 3045 | 7440 |
| 70 | CAA CCC AGT ACC TTG TTA TTG GCA CTG CAA TTT CTT ATT GAT TAA TCA Gln Pro Ser Thr Leu Leu Leu Ala Leu Gln Phe Leu Ile Asp * Ser 3050 3055 3060 | 7488 |
| 75 | GCC AGG AGG AAG GAA ACC TTG GCA CAG TAT CAA CTT GGT ATG TGC ACA Gly Arg Arg Lys Glu Thr Leu Ala Gln Tyr Gln Leu Gly Met Cys Thr 3065 3070 3075 3080 | 7536 |

- 113 -

| | | |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | TGA TGG ATT TAC ACT GGG TGA TTT GGT ACA TAT AAT ACC AAG TCA ATT * Trp Ile Tyr Thr Gly * Phe Gly Thr Tyr Asn Thr Lys Ser Ile 3085 3090 3095 | 7584 |
| 5 | TAC CAA ATG GGG AGA CCA ATA GAG ATG GAG AAA ATC ACA ATC TTA GCT Tyr Gln Met Gly Arg Pro Ile Glu Met Glu Lys Ile Thr Ile Leu Ala 3100 3105 3110 | 7632 |
| 10 | GGA ATT GTG GGG AGG TAA TTC TGA ACT CTC CTT TTT TGA AAT TTT Gly Ile Val Gly Arg * Phe * Thr Leu Leu Phe Phe * Asn Phe 3115 3120 3125 | 7680 |
| 15 | CAT GCT TTA CAT AAT AGT CAA ATG GCT GAC AAA TGT CGT TGT ATG GTT His Ala Leu His Asn Ser Gln Met Ala Asp Lys Cys Arg Cys Met Val 3130 3135 3140 | 7728 |
| 20 | CTC TCT ACC TAA ACC GTT AAG GCA GTA AGA GTT TCC CTA CAA GAT CTC Leu Ser Thr * Thr Val Lys Ala Val Arg Val Ser Leu Gln Asp Leu 3145 3150 3155 3160 | 7776 |
| 25 | TTT GTT CGT ATA ATT GTA TTT TCT AGA GAA AAG TTG CCT TCA ATT TTG Phe Val Arg Ile Ile Val Phe Ser Arg Glu Lys Leu Pro Ser Ile Leu 3165 3170 3175 | 7824 |
| 30 | TGC ACG CGG CAG TAC AGG AAT TGT GGT TAT AAA TAT TGA TAC AGG CTG Cys Thr Arg Gln Tyr Arg Asn Cys Gly Tyr Lys Tyr * Tyr Arg Leu 3180 3185 3190 | 7872 |
| 35 | ACC ATC GTT ACT AAT AGG GGG AAC AAT AAG CAC ATT TTT TTA ATA GCA Thr Ile Val Thr Asn Arg Gly Asn Asn Lys His Ile Phe Leu Ile Ala 3195 3200 3205 | 7920 |
| 40 | AAG GCA TCA CCC TTG TTC CGT TTC CAA TGA AAT CAC AGT ATC CGA ACC Lys Ala Ser Pro Leu Phe Arg Phe Gln * Asn His Ser Ile Arg Thr 3210 3215 3220 | 7968 |
| 45 | ATA AGT TTT ACA AGT ATG CGT AGA GAG AAA TAA AGT ATC AAC CCG GCA Ile Ser Phe Thr Ser Met Arg Arg Glu Lys * Ser Ile Asn Pro Ala 3225 3230 3235 3240 | 8016 |
| 50 | GAA ACA GTT GTT TCA GGC GCA AAG AGA AAA GGA AAC GAT ATG CTC TAT Glu Thr Val Val Ser Gly Ala Lys Arg Lys Gly Asn Asp Met Leu Tyr 3245 3250 3255 | 8064 |
| 55 | TAC ATC AAC CTT TTA GCA TTT AGG GAC GAC CAG CAT CAT CCC ATC TTC Tyr Ile Asn Leu Leu Ala Phe Arg Asp Asp Gln His His Pro Ile Phe 3260 3265 3270 | 8112 |
| 60 | AAT CAA CTG GAG CGA GGT CAC CTC CAA TCT TCT CAG CAG CCT CAG AGT Asn Gln Leu Glu Arg Gly His Leu Gln Ser Ser Gln Gln Pro Gln Ser 3275 3280 3285 | 8160 |
| | GGT GAC CTC CCA AGC AAG TGC ATC AGC ATC CAT CAT CTG GGG GTT GGG Gly Asp Leu Pro Ser Lys Cys Ile Ser Ile His His Leu Gly Val Gly 3290 3295 3300 | 8208 |
| | CAC ATA CCA TGA GCA CAA TCA CCT GAA TTT GAT GAA TTT TCC TCT GTT His Ile Pro * Ala Gln Ser Pro Glu Phe Asp Glu Phe Ser Ser Val 3305 3310 3315 3320 | 8256 |
| | TAC CTT GCA GCA GAC CCC TGC CGT ATA AAT GGT TTT AAA TGA CAG CAT Tyr Leu Ala Ala Asp Pro Cys Arg Ile Asn Gly Phe Lys * Gln His 3325 3330 3335 | 8304 |

- 114 -

| | | |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| | GTT CTT TCA GTT TGA GCA AAA TTT GTG CAA TTG CAA AGA AGC TTT AGA Val Leu Ser Val * Ala Lys Phe Val Gln Leu Gln Arg Ser Phe Arg 3340 3345 3350 | 8352 |
| 5 | ATC ATG TGG AAC ATG CAC TTA CAT TTC ATC TGA CAA TAT AGG AAG GAG Ile Met Trp Asn Met His Leu His Ile * Gln Tyr Arg Lys Glu 3355 3360 3365 | 8400 |
| 10 | AGC CCG ACG TCG CAT GCT CCT CTA GAC TCG AGG AAT TCG CAA GAT TGT Ser Pro Thr Ser His Ala Pro Leu Asp Ser Arg Asn Ser Gln Asp Cys 3370 3375 3380 | 8448 |
| 15 | CTG TCA AAA GAT TGA GGA AGA GGC AGA TGC GCA ATT TCT TTG TTT GTC Leu Ser Lys Asp * Gly Arg Gly Arg Cys Ala Ile Ser Leu Phe Val 3385 3390 3395 3400 | 8496 |
| 20 | TCA TGG TTT CTC AAG TAA GAC TTA TAT CTG ATC TCT TCA ATT TTT GAG Ser Trp Phe Leu Lys * Asp Leu Tyr Leu Ile Ser Ser Ile Phe Glu 3405 3410 3415 | 8544 |
| 25 | ATT GCC TGT TTT TCA CAA TGG CAT ATG TTG TCA GGT GAA ACA TCC AAT Ile Ala Cys Phe Ser Gln Trp His Met Leu Ser Gly Glu Thr Ser Asn 3420 3425 3430 | 8592 |
| 30 | CCC AGT ATT AAT AGA GCC AAC ATG AAG GGA TTG CTT ATC TGA GAT ATC Pro Ser Ile Asn Arg Ala Asn Met Lys Gly Leu Ile * Asp Ile 3435 3440 3445 | 8640 |
| 35 | TGC CAA AGT TGA ATT CTT AGA TTC ACC TTC TTC AGT ATT TCA GAC CTT Cys Gln Ser * Ile Leu Arg Phe Thr Phe Ser Ile Ser Asp Leu 3450 3455 3460 | 8688 |
| 40 | CTA AGC ATT TTC ATT TTT TTC AAT TGT TAG GGA GTT CCA ATG TTT Leu Ser Ile Phe Ile Phe Phe Asn Cys * Gly Val Pro Met Phe 3465 3470 3475 3480 | 8736 |
| 45 | TAC ATG GGC GAT GAA TAT GGC CAC ACA AAA GGG GGC AAC AAC AAT ACA Tyr Met Gly Asp Glu Tyr Gly His Thr Lys Gly Gly Asn Asn Asn Thr 3485 3490 3495 | 8784 |
| 50 | TAC TGC CAT GAT TCT TAT GTC AGT ACA ATT TGG TCA CAT ATT GTT GTT Tyr Cys His Asp Ser Tyr Val Ser Thr Ile Trp Ser His Ile Val Val 3500 3505 3510 | 8832 |
| 55 | CTA AGT AAC TAT CTT CAA ATC TTT GCA TTC ATC CGT CAT GGC TCT TCT Leu Ser Asn Tyr Leu Gln Ile Phe Ala Phe Ile Arg His Gly Ser Ser 3515 3520 3525 | 8880 |
| 60 | GTA GGT CAA TTA TTT TCG CTG GGA TAA AAA AGA ACA ATA CTC TGA CTT Val Gly Gln Leu Phe Ser Leu Gly * Lys Arg Thr Ile Leu * Leu 3530 3535 3540 | 8928 |
| | GCA AAG ATT CTG CTG CCT CAT GAC CAA ATT CCG CAA GTA AGT ATT CCG Ala Lys Ile Leu Leu Pro His Asp Gln Ile Pro Gln Val Ser Ile Pro 3545 3550 3555 3560 | 8976 |
| | TTG AAT AAT TTC TGT GTA GAA CCA CTG AAG GTG CCT CCA AAC GCT AAG Leu Asn Asn Phe Cys Val Glu Pro Leu Lys Val Pro Pro Asn Ala Lys 3565 3570 3575 | 9024 |
| | CGA GCA AGG TCA ATT TCA CAC CCT AAT CAA GTT GGT GTT GTC TAT TTG Arg Ala Arg Ser Ile Ser His Pro Asn Gln Val Gly Val Val Tyr Leu 3580 3585 3590 | 9072 |

- 115 -

| | | |
|----|------------------------------------------------------------------------------------------------------------------------------------|------|
| | TGT ATT TGA TCT GCT GCA CTG TAG GGA GTG CGA GGG TCT TGG CCT TGA Cys Ile * Ser Ala Ala Leu * Gly Val Arg Gly Ser Trp Pro * | 9120 |
| | 3595 3600 3605 | |
| 5 | GGA CTT TCC AAC GGC CGA ACG GCT GCA GTG GCA TGG TCA TCA GCC TGG Gly Leu Ser Asn Gly Arg Thr Ala Ala Val Ala Trp Ser Ser Ala Trp | 9168 |
| | 3610 3615 3620 | |
| 10 | GAA GCC TGA TTG GTC TGA GAA TAG CCG ATT CGT TGC CTT TTC CAT GGT Glu Ala * Leu Val * Glu * Pro Ile Arg Cys Leu Phe His Gly | 9216 |
| | 3625 3630 3635 3640 | |
| 15 | ACA CAT ATA GTT CTG ACA CTT CAC TAT AGT TGT TTT AAA AAA GAA AAT Thr His Ile Val Leu Thr Leu His Tyr Ser Cys Phe Lys Lys Glu Asn | 9264 |
| | 3645 3650 3655 | |
| | TTA ACT CAA AAG TAA ATT ATG GAG A Leu Thr Gln Lys * Ile Met Glu | 9289 |
| 20 | 3660 | |

DRAFTING & DESIGN